

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 09:56:59 ; Search time 34 Seconds  
(without alignments)  
3248.427 Million cell updates/sec

Title: US-10-091-085-7  
Perfect score: 2252  
Sequence: 1 MATSWGTVPFMLVVCVCSA.....ETGWALGATFLLQLSGLISH 428

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTRMBL\_23:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phase:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_rvirus:\*
- 16: sp\_bacteriaph:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	2221	98.6	428	4 Q96RX0	Q96RX0 homo sapien
2	2089	92.8	407	4 Q8WUB3	Q8WUB3 homo sapien
3	1975.5	87.7	427	11 Q8CD29	Q8CD29 mus musculus
4	1969.5	87.5	427	11 Q8BR23	Q8BR23 mus musculus
5	993	44.1	483	4 Q8TAS7	Q8TAS7 homo sapien
6	993	44.1	503	4 Q8N3H3	Q8N3H3 homo sapien
7	693.5	30.8	461	5 Q76268	Q76268 drosophila
8	693.5	30.8	464	5 Q9VQ18	Q9VQ18 drosophila
9	616.5	27.4	479	5 Q9XU84	Q9XU84 caenorhabdi
10	590	26.2	278	11 Q8CHZ3	Q8CHZ3 mus musculus
11	535.5	23.8	556	3 Q9UT35	Q9UT35 schizosacch
12	513.5	22.8	489	10 Q8H7L6	Q8H7L6 oryza sativ
13	513.5	22.8	599	3 Q8TGH6	Q8TGH6 candida alb
14	510	22.6	522	3 Q9HEM6	Q9HEM6 kluyveromyc
15	507.5	22.5	467	10 Q9SPM7	Q9SPM7 dolichos bi
16	502.5	22.3	455	10 Q9SPM6	Q9SPM6 medicago sa

17	493	21.9	462	10 Q9XFC9	Q9XFC9 dolichos bi
18	492	21.8	463	10 Q9FVC3	Q9FVC3 glycine soj
19	486.5	21.6	456	10 Q9SPM8	Q9SPM8 locus japon
20	484	21.5	455	10 Q9FEA6	Q9FEA6 pieum sativ
21	484	21.5	472	10 Q8L704	Q8L704 arabidopsis
22	483	21.4	472	10 Q9SPM5	Q9SPM5 arabidopsis
23	481	21.4	455	10 Q8GTB1	Q8GTB1 pieum sativ
24	480	21.3	472	10 Q8M7B3	Q8M7B3 arabidopsis
25	479	21.3	471	10 Q9SQG2	Q9SQG2 arabidopsis
26	476	21.1	447	10 Q9SLV4	Q9SLV4 pieum sativ
27	471	20.9	467	10 Q8RVT6	Q8RVT6 pieum sativ
28	470	20.9	468	10 Q8FVC2	Q8FVC2 glycine soj
29	467	20.7	407	10 Q9AVN8	Q9AVN8 pieum sativ
30	467	20.7	455	10 Q8RVU0	Q8RVU0 pieum sativ
31	462	20.5	447	10 Q9FU11	Q9FU11 pieum sativ
32	457.5	20.3	558	3 Q8TGG8	Q8TGG8 aspergillus
33	457	20.3	466	10 Q9AU15	Q9AU15 medicago tr
34	444	19.7	455	10 Q8RVT9	Q8RVT9 pieum sativ
35	438	19.4	473	10 Q8RVT8	Q8RVT8 pieum sativ
36	433.5	19.2	555	10 Q94E22	Q94E22 arabidopsis
37	432.5	19.2	271	5 Q8IPZ6	Q8IPZ6 drosophila
38	431.5	19.2	556	3 Q9C2M0	Q9C2M0 neurospora
39	430.5	19.1	555	10 Q80612	Q80612 arabidopsis
40	421.5	18.7	224	11 Q8CDB6	Q8CDB6 mus musculus
41	403	17.9	493	13 Q90X66	Q90X66 gallus gall
42	398	17.7	483	10 Q9XI62	Q9XI62 arabidopsis
43	385	17.1	539	11 Q8CDV7	Q8CDV7 mus musculus
44	382	17.0	510	11 Q921Q6	Q921Q6 mus musculus
45	374	16.6	690	5 Q9BHV5	Q9BHV5 leishmania

## ALIGNMENTS

## RESULT 1

Q96RX0	PRELIMINARY;	PRT;	428 AA.
ID	Q96RX0		
AC	Q96RX0;		
DT	01-DEC-2001 (TrEMBLrel. 19, Created)		
DT	01-DEC-2001 (TrEMBLrel. 19, Last sequence update)		
DT	01-MAR-2002 (TrEMBLrel. 20, Last annotation update)		
DE	Pcph proto-oncogene protein.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=20173601; PubMed=10708485;		
RA	Recio J.A., Zambrano N., Pena, Ld, Reig J.A., Rhoads A., Rouzaut A.,		
RA	Notario V.;		
RT	"The human PCPH proto-oncogene: cDNA identification, primary		
RT	structure, chromosomal mapping, and expression in normal and tumor		
RT	cells.";		
RL	Mol. Carcinog. 27:229-236(2000).		
DR	EMBL; AF136572; AAK82950.1; -		
DR	InterPro; IPR000407; GDAL_CD39_NTPase.		
DR	Pfam: PF01150; GDAL_CD39_1.		
SQ	SEQUENCE 428 AA; 47431 MW; P2C4F7DE650A44F6 CRC64;		
Query Match	98.6%;	Score 2221;	DB 4; Length 428;
Best Local Similarity	98.8%;	Pred. No. 3.3e-183;	
Matches 423;	Conservative	0; Mismatches 15;	Indels 0; Gaps 0;
Qy	1 MATSWGTVPFMLVVCVCSA	VSHRNQOTWFEGLFLSSMCPINVSASTLYGIMFDAGSTGT	60
Db	1 MATSWGTVPFMLVVCVCSA	VSHRNQOTWFEGLFLSSMCPINVSASTLYGIMFDAGSTGT	60
Qy	61 RIHVYTFVQMPQLPILEGEVDSVKPGLSAFVDOPKQGAETVQGLLELVAKDSIPRSHW	120	
Db	61 RIHVYTFVQMPQLPILEGEVDSVKPGLSAFVDOPKQGAETVQGLLELVAKDSIPRSHW	120	
Qy	121 KKTTPVVLKATAGURLLPEHKAKALLFEVKBIFRKSPPFLVPKGSVIMTQDEGIFAWTV	180	

```
Db 121 KKTFWLKAATAGRLLEPHKAKALLFEVKEIFRKSPLVPKGSVIMDGSDEGLAWTV 180
Qy 181 NFLTGLQHGRQETVGTLDLGGASTQITFLPQPEKTLQTPRGYLTSEFNFSTYKLYTH 240
Db 181 NFLTGLQHGRQETVGTLDLGGASTQITFLPQPEKTLQTPRGYLTSEFNFSTYKLYTH 240
Qy 241 SYLGFGLKAARLATLGALETGTDTGHTFRSACLPRWLEAEWIFGVKYQYGGNQEVEVF 300
Db 241 SYLGFGLKAARLATLGALETGTDTGHTFRSACLPRWLEAEWIFGVKYQYGGNQEVEVF 300
Qy 301 EPCYAEVLVRVVRGKLHQPVEEVQSGSFYAFSYYVDRAVDTMDIDYKGGILKVEDFERKAR 360
Db 301 EPCYAEVLVRVVRGKLHQPVEEVQSGSFYAFSYYVDRAVDTMDIDYKGGILKVEDFERKAR 360
Qy 361 EVCNLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQ 400
Db 361 EVCNLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQ 400
Qy 421 LQSLGISH 428
Db 421 LQSLGISH 428

RESULT 2
Q8WUB3 PRELIMINARY; PRT; 407 AA.
AC O8WUB3;
DT 01-MAR-2002 (TReMBLrel. 20, Created)
DT 01-MAR-2002 (TReMBLrel. 20, last sequence update)
DE Similar to ectonucleoside triphosphate diphosphohydrolase 5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OK NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Colon;
RA Strauberg R.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC020966; AAH20966.1; -.
DR InterPro; IPR000407; GDAI_CD39_NTPase.
DR Pfam; PF01150; GDAI_CD39; 1.
KW Hydrolase.
SQ SEQUENCE 407 AA; 45336 MW; D92A5F7DC9EC9E5B CRC64;

Query Match 92.8%; Score 2089; DB 4; Length 407;
Best Local Similarity 99.2%; Pred. No. 7.7e-172;
Matches 397; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MATSWGTVFFMLVVCVCSAVSHRNQQTWFEIGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
Db 1 MATSWGTVFFMLVVCVCSAVSHRNQQTWFEIGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
Qy 61 RIHYTTFVQKMPGQLPILEGEVDFSVKPGLSAFVDQPKQGAETVQGLLEAKDSIPRSHW 120
Db 61 RIHYTTFVQKMPGQLPILEGEVDFSVKPGLSAFVDQPKQGAETVQGLLEAKDSIPRSHW 120
Qy 121 KKTFWLKAATAGRLLEPHKAKALLFEVKEIFRKSPLVPKGSVIMTQDEGIFAWTV 180
Db 121 KKTFWLKAATAGRLLEPHKAKALLFEVKEIFRKSPLVPKGSVIMTQDEGIFAWTV 180
Qy 181 NFLTGLQHGRQETVGTLDLGGASTQITFLPQPEKTLQTPRGYLTSEFNFSTYKLYTH 240
Db 181 NFLTGLQHGRQETVGTLDLGGASTQITFLPQPEKTLQTPRGYLTSEFNFSTYKLYTH 240
Qy 241 SYLGFGLKAARLATLGALETGTDTGHTFRSACLPRWLEAEWIFGVKYQYGGNQEVEVF 300
Db 241 SYLGFGLKAARLATLGALETGTDTGHTFRSACLPRWLEAEWIFGVKYQYGGNQEVEVF 300
Qy 301 EPCYAEVLVRVVRGKLHQPVEEVQSGSFYAFSYYVDRAVDTMDIDYKGGILKVEDFERKAR 360
Db 301 EPCYAEVLVRVVRGKLHQPVEEVQSGSFYAFSYYVDRAVDTMDIDYKGGILKVEDFERKAR 360
Qy 361 EVCNLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQ 400
Db 361 EVCNLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQ 400
Qy 421 LQSLGISH 428
Db 421 LQSLGISH 428

RESULT 3
Q8CD29 PRELIMINARY; PRT; 427 AA.
AC Q8CD29;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, last sequence update)
DE Ectonucleoside triphosphate diphosphohydrolase 5.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OK NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
DR EMBL; AK031581; BAC27461.1; -.
SQ SEQUENCE 427 AA; 47101 MW; 6E3773C842B58477 CRC64;

Query Match 87.7%; Score 1975.5; DB 11; Length 427;
Best Local Similarity 87.4%; Pred. No. 5.2e-162;
Matches 373; Conservative 25; Mismatches 28; Indels 1; Gaps 1;

Qy 1 MATSWGTVFFMLVVCVCSAVSHRNQQTWFEIGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
Db 1 MATSWGAV-FMLIIACVGVSTVFYREQQTWFEIGIFLSSMCPINVSAGTFYIGIMFDAGSTGT 59
Qy 61 RIHYTTFVQKMPGQLPILEGEVDFSVKPGLSAFVDQPKQGAETVQGLLEAKDSIPRSHW 120
Db 60 RIHYTTFVQKTAGQLPFLGEIGIFDSVKPGLSAFVDQPKQGAETVQGLLEAKDSIPRSHW 119
Qy 121 KKTFWLKAATAGRLLEPHKAKALLFEVKEIFRKSPLVPKGSVIMTQDEGIFAWTV 180
Db 120 ERTFVVLKATAGLRLPEQKAQALLLEVEBEIFKNSPFLVDPGVSIMDGSYEGILAWTV 179
Qy 181 NFLTGLQHGRQETVGTLDLGGASTQITFLPQPEKTLQTPRGYLTSEFNFSTYKLYTH 240
Db 180 NFLTGLQHGRQETVGTLDLGGASTQITFLPQPEKTLQTPRGYLTSEFNFSTYKLYTH 239
Qy 241 SYLGFGLKAARLATLGALETGTDTGHTFRSACLPRWLEAEWIFGVKYQYGGNQEVEVF 300
Db 240 SYLGFGLKAARLATLGALEAKGTDGHTFRSACLPRWLEAEWIFGVKYQYGGNQEVEVF 299
Qy 301 EPCYAEVLVRVVRGKLHQPVEEVQSGSFYAFSYYVDRAVDTMDIDYKGGILKVEDFERKAR 360
Db 300 EPCYAEVLVRVVRGKLHQPVEEVQSGSFYAFSYYVDRAADTHLIDYKGGVLKVEDFERKAR 359
Qy 361 EVCNLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQ 400
Db 360 EVCNLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQ 400
Qy 421 LQSLGISH 427
Db 420 LQSLGIT 426

RESULT 4
Q8BR23 PRELIMINARY; PRT; 427 AA.
AC Q8BR23;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
```

```
Db 301 EPCYAEVLVRVVRGKLHQPVEEVQSGSFYAFSYYVDRAVDTMDIDYKGGILKVEDFERKAR 360
Qy 361 EVCNLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQ 400
Db 361 EVCNLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQ 400

RESULT 3
Q8CD29 PRELIMINARY; PRT; 427 AA.
AC Q8CD29;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, last sequence update)
DE Ectonucleoside triphosphate diphosphohydrolase 5.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OK NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
DR EMBL; AK031581; BAC27461.1; -.
SQ SEQUENCE 427 AA; 47101 MW; 6E3773C842B58477 CRC64;

Query Match 87.7%; Score 1975.5; DB 11; Length 427;
Best Local Similarity 87.4%; Pred. No. 5.2e-162;
Matches 373; Conservative 25; Mismatches 28; Indels 1; Gaps 1;

Qy 1 MATSWGTVFFMLVVCVCSAVSHRNQQTWFEIGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
Db 1 MATSWGAV-FMLIIACVGVSTVFYREQQTWFEIGIFLSSMCPINVSAGTFYIGIMFDAGSTGT 59
Qy 61 RIHYTTFVQKMPGQLPILEGEVDFSVKPGLSAFVDQPKQGAETVQGLLEAKDSIPRSHW 120
Db 60 RIHYTTFVQKTAGQLPFLGEIGIFDSVKPGLSAFVDQPKQGAETVQGLLEAKDSIPRSHW 119
Qy 121 KKTFWLKAATAGRLLEPHKAKALLFEVKEIFRKSPLVPKGSVIMTQDEGIFAWTV 180
Db 120 ERTFVVLKATAGLRLPEQKAQALLLEVEBEIFKNSPFLVDPGVSIMDGSYEGILAWTV 179
Qy 181 NFLTGLQHGRQETVGTLDLGGASTQITFLPQPEKTLQTPRGYLTSEFNFSTYKLYTH 240
Db 180 NFLTGLQHGRQETVGTLDLGGASTQITFLPQPEKTLQTPRGYLTSEFNFSTYKLYTH 239
Qy 241 SYLGFGLKAARLATLGALETGTDTGHTFRSACLPRWLEAEWIFGVKYQYGGNQEVEVF 300
Db 240 SYLGFGLKAARLATLGALEAKGTDGHTFRSACLPRWLEAEWIFGVKYQYGGNQEVEVF 299
Qy 301 EPCYAEVLVRVVRGKLHQPVEEVQSGSFYAFSYYVDRAVDTMDIDYKGGILKVEDFERKAR 360
Db 300 EPCYAEVLVRVVRGKLHQPVEEVQSGSFYAFSYYVDRAADTHLIDYKGGVLKVEDFERKAR 359
Qy 361 EVCNLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQ 400
Db 360 EVCNLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQ 400
Qy 421 LQSLGISH 427
Db 420 LQSLGIT 426

RESULT 4
Q8BR23 PRELIMINARY; PRT; 427 AA.
AC Q8BR23;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
```



```
Qy 97 PQGAETVQGLLEVAKDSIPRSHWKKTPVVLKATAGLRLPEHAKALLFEVKEIFRKPSP 156
Db 168 VEKSAQGIITRELLDVAQDIPDFWKATPLVLKATAGLRLPEHAKALLFEVKEIFRKPSP 227
Qy 157 FLVPGKSGVINTGDEGIFAWVTNFLTGLQHGHRQETVGTLDLGGASTQITLPQREKT 216
Db 228 FLVGDDCVSINMGDEGSAWITINFLTGLQHGHRQETVGTLDLGGASTQITLPQREKT 287
Qy 217 LEQTPRGVLTSEFNFNSTYKLYTHSYLGFGLKAARLATLGALETE-GTDTGHTFSACLPR 275
Db 288 LQASPPGYLTALRNFNTYKLYSYLGLGLMSARLAILGGVEQPAKDGKELVSPCLSP 347
Qy 276 WLEAEWIFGGVYQYQGNQGEVGFEPYAEVLVRVGRKHLQHPQEVQSGFYAFSYYYDR 335
Db 348 SFKGEWEHAEVTVRVSGQAAASLHELCAARVSEVLQNRVHRTVEVKHVDVYAFSYYYDL 407
Qy 336 AVDTMDIDYKGGILKVEDFERKAREVCDNLNENTSGSPFLCMLDLSITALLKDGFGFAD 395
Db 408 AAGVGLIDAEGKGLVVGDFEIAAKYVCRITLQFQSSPFSCMDLTVYVLLQLQ-EGFPR 466
Qy 396 STVLQTLKKNVNIETGALGATFHLQSL 424
Db 467 SKVLKLRKIDNVETSWALGAIHFYIDSL 495
RESULT 7
ID 076268 PRELIMINARY; PRT; 461 AA.
AC 076268;
DT 01-NOV-1998 (TrEMBLrel. 08, Created)
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE NTPASE protein (L111641p).
GN NTPASE OR CG3059.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celnikner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champagne M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abriel J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Folsler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusekern D.R., Pacleb J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
```

```
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhou S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RA "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RP SEQUENCE FROM N.A.
RX MEDLINE=98341119; PubMed=96764430;
RA Chadwick B.P., Frischauf A.-M.;
RT "The CD39-like gene family: identification of three new human members
RT (CD39L2, CD39L3, and CD39L4), their murine homologues, and a member of
RT the gene family from Drosophila melanogaster.";
RL Genomics 50:357-367(1998).
RP SEQUENCE FROM N.A.
RC STRAIN=Berkeley;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
RA Nunoo J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,
RA Yu C., Lewis S.E., Rubin G.M., Celnikner S.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE003581; AAF51182.1; -
DR EMBL; AF041048; AAC39133.1; -
DR EMBL; AY061134; AAL28682.1; -
DR FlyBase; FBgn0024947; NTPase.
DR InterPro; IPR000407; GDAL_CD39_NTPase.
DR Pfam; PF01150; GDAL_CD39; 1.
SQ SEQUENCE 461 AA; 50845 MW; 27D00321F91A9DD1 CRC64;
Query Match 30.8%; Score 693.5; DB 5; Length 461;
Best Local Similarity 39.1%; Pred. No. 3.2e-51;
Matches 156; Conservative 73; Mismatches 123; Indels 47; Gaps 12;
Qy 49 YGIMFDAGSTGTRIHVYTFVQKMPGQLPILGEVDFSVKPLGSFAFVDQPKQGAETVQGLL 108
Db 79 YAAIDAGSTGSRVLAYKFNRSFIDNKLVLVEYELFKERKPLGSFADNPAEGAHSIKLL 138
Qy 109 EVAKDSIPRSHWKKTPVVLKATAGLRLPEHAKALLFEVKEIFRKSFFLVPGKSGVIMT 168
Db 139 DEARAFIPKEHWSSTPLVLKATAGLRLPASKAENILNAVRLFAKSEFSVDMDAVEIMD 198
Qy 169 GQDEGIFAWVTNFLTGLQHGHRQETVGTLDLGGASTQITF-----LPQEKTLQTP 221
Db 199 GTDEGIFSWFTNFLTGLRUKTNQ--AAALDLGGSGTQVTFSTPDQDPVPYDKYMHV- 255
Qy 222 RGYLTSEFNFNSTYKLYTHSYLGFGLKAARLATLGALETEG--TDGHTFSACL-PRWLE 278
Db 256 ---VTSKKIN---VFTHSYLGLGLMAARHAFV---THGYKKEDTVLESVCNPIAN 304
Qy 279 AEWIFGGVYQYQGNQGE-----VGPEPCVAEVLVRVGRKGL-----HQPBEVQSGSF 326
Db 305 RTWTYGNVQYKSGKNGKSSARQPIVDFDAC-----LELVKSMPLVLPKPKFTLKQHAV 360
Qy 327 YAESYYDRAVDMDIDYKGGILKVEDFERKAREVC--DNLENFTSGSPFLCMLDLSYIT 384
Db 361 AAFSYFPERAIESGLVDPLAGGTTVEAYRKKAEICAI PNDE-----QPFMCPDLTFTIS 415
Qy 385 ALIKDGFADSTVQLTKKNVNIETGALGATFHLQSL 423
Db 416 TLLREGFLNDGKKIKLYKIDGHEISWALGCAYNVLTIS 454
RESULT 8
Qy 09VQ18 PRELIMINARY; PRT; 464 AA.
ID 09VQ18
AC 09VQ18;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
```



```

DE NTPAS2 protein.
GN NTPASE OR CG3059.
OS Drosophila melanogaster (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY.
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Holt G., Nelson C.R., Miklos G.L.G.,
RA April J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballwe R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bertram B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karp G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacieb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL; AE003581; AAF51181.1; -.
DR FlyBase; FBgn0024947; NTPase.
DR InterPro; IPR000407; GDAL_CD39_NTPase.
DR Pfam; PF01150; GDAL_CD39; 1.
SQ SEQUENCE 464 AA; 5119 MW; 71D057ABE5AE613D CRC64;

Query Match 30.8%; Score 693.5; DB 5; Length 464;
Best Local Similarity 39.1%; Pred. No. 3.2e-51;
Matches 156; Conservative 73; Mismatches 123; Indels 47; Gaps 12;

QY 49 YGIMFDAGSTGTRIHYVTYVQKMPGQQLPILEGVEFDSVKPGLSAFVDQPKQGAETVOGLL 108
DB 82 YAAIIDAGSTGSRVLYAKFNRSFDINKLVLYEELFKERKPGLSFADNPAEGAHSKILL 141
QY 109 EVAKDSIPRSHWKTTPVLKATAGLRLLPEHKAALLFEYKEIFRKSPLVPKGSYSIMT 168
DB 142 DEARAFIPKEHWSSTPLVLKATAGLRLLPASKAENILNAVDRULFAKSEFSDMDAVEIMD 201
QY 169 GODEGIFAWTVNFLTQLGHQROETVGTLDGGASTQITFLPQFKEKLEOTPRGYLTSFE 229
DB 202 GTDEGIFSWTVNFLTGLRLSEKTNQ--AAALDLGGSTQVTFSPDPQVFPYDKYHVEV- 258
QY 222 RGYLTSEMFNSTYKLYTHSYLGLKAARLATIAGALETEG--TDGHTFRSACL-PRWLE 278

```

```

259 ----VTSSKKIN-----VFTHSYLGLGLMAARHAFV-----THGYKKEDTVLESVCNPLIAN 307
279 AEWIFGGVKYQYGNQEGE-----VGPEPCYAEVLRVVRGKL-----HQPEEVQRGSP 326
308 RTWTYGYQYQYKVGKENGKSSAEQPIVDPDAC---LELVKSKVMPLVKPKPPTLKQHAV 363
327 YAFSYYYDRAVDTMDIDYEKGILKVEDFERKAREVC--DNLENFTSGSPFLCMDLSYIT 384
364 AAFSYFFERAIESGLVPLAGGETTVYAYRKAQCAIPNDE-----QFFMCFDLTTFIS 418
385 ALLKDGFGFADSTVLQLTKKVNIETGMAWGATFHLQOS 423
419 TLRREGFLNGDKKIKLYKKIDGHEISWALGAYNVLTS 457

RESULT 9
Q9XU84 PRELIMINARY; PRT; 479 AA.
AC Q9XU84;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE K08H10.4 protein.
GN K08H10.4.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidae;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA Gardner A.E.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL; Z83113; CAB05544.1; -.
DR WormPep; K08H10.4; CE18877.
DR InterPro; IPR002034; AIPM/HcIt_synth.
DR InterPro; IPR000407; GDAL_CD39_NTPase.
DR Pfam; PF01150; GDAL_CD39; 1.
DR PROSITE; PS00815; AIPM HOMOCIT SYNTH 1; 1.
SQ SEQUENCE 479 AA; 5384 MW; 7EDC02A9D54A48ED CRC64;

Query Match 27.4%; Score 616.5; DB 5; Length 479;
Best Local Similarity 34.1%; Pred. No. 1.5e-44;
Matches 149; Conservative 82; Mismatches 167; Indels 39; Gaps 16;

QY 9 FFMVLVYS--CVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGTRIHYVT 66
DB 6 FSILLISFSLSVVTKTQY-WCHGDGVLN-----NQHTCRPTTIVIDAGSTGTRLHLX 60
QY 67 FVQK-----MPGQLPILEGVEFDSVKPGLSAFVDQPKQGAETVOGLLEVAKDSIPRSH 119
DB 61 FTHDPAIASHGMPFK--VEKEIFQVQKPGLSFSAKSPSSAASDLPLQARKEVPHFM 117
QY 120 WKTPVVLKATAGLRLLPEHKAALLFEYKEIFRKSPLVPKGSYSIMTQDEGIFAWV 178
DB 118 WEKTPITLKATAGLRLLPGDMADDILESVEERIFNSGFFAAPFPAVNVMPGSEGVYSWF 177
QY 179 TVNFLTGLH-----GHR---QETVGTLDGGASTQITFLPQFKEKLEOTPRGYLTSFE 229
DB 178 TMLILETFTDDEPTVGHKPAHRSAVAFDLGGSTQLTWPNNFAVSEHV-GYERDID 236
QY 230 MFNSTYKLYTHSYLGLKAARLATIAGALETEG--TDGHTFRSACLPRWLE-AEWIFGV 286
DB 237 FFGHRLFTSHPLGNGLAARLNIL-QLETDNEIESTHQLITSCMPEGYQLEWEY-AL 294
QY 287 KYQYGNQEGVEGPEPCYAEVLRVVR-GKHQPEEVQRGSFYAFSYYYDRAVDTMDIYE 345

```

Db 295 KF-WNINGSSHSPESCVCYGTNNFVSSEIMHLRELKSGSPVYLFPSYFFDRALNSGLVKGN 353

Qy 346 KGGILKVEDPERKAREVC-----DNLENFTSGPFLMDLSVITALLKDGFGFADSTVLQL 401

Db 354 EGGKIELRQFKEAREIACREKTEIDGSHWMPQCCLDIYIYSLLRDGGYQFEDNQPLVL 413

Qy 402 TKKVNNIETGWALGATF 418

Db 414 AKIKGMEVSWGQGLAF 430

RESULT 10

Q8CHZ3 PRELIMINARY; PRT; 278 AA.

ID Q8CHZ3

AC Q8CHZ3

DT 01-MAR-2003 (Tremblrel. 23, Created)

DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)

DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)

DE Similar to ectonucleoside triphosphate diphosphohydrolase 6.

OS Mus musculus (Mouse)

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Mammary gland;

RA Strausberg R.;

RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC038126; AAH38126.1; -

KW Hydrolase.

SQ SEQUENCE 278 AA; 30130 MW; 75A92DD1AC76297F CRC64;

Query Match 26.2%; Score 590; DB 11; Length 278;

Best Local Similarity 59.08; Pred. No. 1.3e-42;

Matches 119; Conservative 29; Mismatches 51; Indels 6; Gaps 2;

Qy 49 YGIMFDAGSTGTRIHVTVFQKMPGQLPILEGEVDSVKPGLSAFVDQPKQAGATVQGLL 108

Db 74 YGIMFDAGSTGTRIHVTVFQKMPGQLPILEGEVDSVKPGLSAFVDQPKQAGATVQGLL 132

Qy 109 EVAKDSIPRSHWKTTPVLKATAGLRLPEHAKALLFEVKEIFRKSPELVKPGSVSINT 168

Db 133 NVAKQHIIYPDFWKATPLVLKATAGLRLPEHAKALLFEVKEIFRKSPELVKPGSVSINT 192

Qy 169 GQDEGIFAWTVNFTLQGLHGHQETVGLDGGASTQITFLPOFEKTLBOTPRGYLTSF 228

Db 193 GTDEGSANITVNFLLTSLKTPGSSSVGMLDGGSTQITFLPRVECTLQASPPGHLTAL 252

Qy 229 EMFNSTYKLYTHSVGLGKLAARLA 253

Db 253 QMFNRYKLYSRWV-----CSRLA 272

RESULT 11

Q9UT35 PRELIMINARY; PRT; 556 AA.

ID Q9UT35

AC Q9UT35

DT 01-MAY-2000 (Tremblrel. 13, Created)

DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)

DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)

DE Putative guanosine-diphosphatase (Guanosine diphosphatase).

GN SPAC824.08 OR GDPI.

OS Schizosaccharomyces pombe (Fission yeast).

OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;

OC Schizosaccharomycetales; Schizosaccharomycetaceae;

OC Schizosaccharomycetes.

OX NCBI\_TaxID=4896;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=972h-;

RA Barrell B.G.; Rajandream M.A.; Quail M.; Seegar K.; Harris D.;

RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RA Sanchez R., Franco A., Notario V., Gacto M., Cansado J.;

RT "Characterization of a guanosine diphosphatase gene from

RT Schizosaccharomyces pombe.";

RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; ALI21741; CAB57338.1; -

DR EMBL; AF465240; AAL69974.1; -

DR GeneDB SPombe; SPAC824.08; -

DR InterPro; IPR000407; GDAL\_CD39\_NTPase.

DR Pfam; PF01150; GDAL\_CD39\_1

DR PROSITE; PS01238; GDAL\_CD39\_NTPASE; 1.

SQ SEQUENCE 556 AA; 61588 MW; 1D811E3D6A6BB85 CRC64;

Query Match 23.8%; Score 535.5; DB 3; Length 556;

Best Local Similarity 34.1%; Pred. No. 1.9e-37;

Matches 143; Conservative 61; Mismatches 158; Indels 57; Gaps 13;

Qy 49 YGIMFDAGSTGTRIHVTVFQKMPGQLPILEGEVDSVKPGLSAFVDQPKQAGATVQGLL 108

Db 134 YVLMIDAGSTGSRVHVYQFNCPNS--PKLEEEFFKMIPEGLSSFAGDPEGAAASLDPLL 191

Qy 109 EVAKDSIPRSHWKTTPVLKATAGLRLPEHAKALLFEVKEIFRKSPELVKPGSVSINT 167

Db 192 DYAMENVPEYRRCSPIAVKATAGLRLTGESEKAILKSVRQHLNDYPPPIVDGVISIL 251

Qy 168 TGQDEGIFAWTVNFTLQGLHGHQETVGLDGGASTQITFLPOFEKTLBOTPRG--- 223

Db 252 EGSMEGIYAWITINYLTLGKKAHTSTVAVMDLGGASTQITFLPRFASGESIVDGDHK 311

Qy 224 YLTSFEMFNSTYKLYTHSVGLGKLAARL-----ATLGALETGTDGHTFRSAC 272

Db 312 YVLDYN--GEQYELYQSHLGYGLKEARKLIHFLVNNAEALKESLELLG-DSTSIHPC 368

Qy 273 L-----PWLAEWIFEGVKYQYGGNORGEVGFPCYAEVLRVVR---GKLLHOP 318

Db 369 LHLNASLTHPDSKSEASEVVFVPSLAHLSLQCRGIAEKALYKDKKCPVRPCSFNGVHQP 428

Qy 319 ---EEVQRGSFAFSYVYDRAVDTMDIYKGGILKVEDPERKAREVC----- 363

Db 429 KFTETTTDPIYLSIFYDR----MISLGMPTFTIEDMKYLANSCSGPTTYQDAFSL 483

Qy 364 -DNLENFTSGPFLMDLSVITALLKDGFGFADSTVLQTKKVNNIETGWALGATFHL 421

Db 484 TDALKELKE-EPWECILDYNMISLLSVGYEIPNNQLHTAKKIDNKELGCLGASLSML 541

RESULT 12

Q8H7L6 PRELIMINARY; PRT; 489 AA.

ID Q8H7L6

AC Q8H7L6

DT 01-MAR-2003 (Tremblrel. 23, Created)

DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)

DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)

DE Putative apyrase.

GN OSJNB0014110.10.

OS Oryza sativa (japonica cultivar-group).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

OC Ehrhartoideae; Oryzoideae; Oryza.

OX NCBI\_TaxID=39947;

RN [1]

RP SEQUENCE FROM N.A.

RA Wing R.A., Yu Y., Soderlund C., Kim H.-R., Rambo T., Saski C.,

RA Currie J., Collura K.;

RT "Rice Genomic Sequence."

RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AC126222; AAN65004.1; -

SQ SEQUENCE 489 AA; 52799 MW; 0BF6BF154488D38E CRC64;

Query Match 22.8%; Score 513.5; DB 10; Length 489;

Best Local Similarity 33.7%; Pred. No. 1.2e-35;

Matches 142; Conservative 70; Mismatches 150; Indels 59; Gaps 16;



Search completed: January 29, 2004, 10:05:31  
Job time : 35 secs

```
QY 309 RVRGKHLQHPVEV-----QRGSFYAFSYYYDRAVDTMIDVEKGGI-----LKVEDFERKARE 361
Db 389 ----NGIHQPSLVHTFKETSLDYVFSFYDRT-----QPLGLPLSFTLQELQDLART 436
QY 362 VCDNLENFTS-----GS-----PFLCMDLSYITALLKDGFGFADSTVLQLTKVNNI 408
Db 437 VCNGEEVWESVFSGIEGSLSELSKEPQWCLDNLFQVSLHTGYDIPLQRELRTAKTIANN 496
QY 409 ETGWALGATPHLLOS 423
Db 497 ELGWCLGASLPLES 511

RESULT 15
Q9SPW7
ID Q9SPM7 PRELIMINARY; PRT; 467 AA.
AC Q9SPM7;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Apyrase.
OS Dolichos biflorus (Horse gram).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Dolichos.
OX NCBI_TaxID=3840;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99444909; PubMed=10517321;
RA Roberts N.J., Brigham J., Wu B., Murphy J.B., Volpin H.,
RA Phillips D.A., Etzler M.E.;
RT "A Nod factor-binding lectin is a member of a distinct class of
RT apyrases that may be unique to the legumes.";
RL Mol. Gen. Genet. 262:261-267(1999).
DR EMBL; AF156781; AA00610.1; -
DR InterPro; IPR000407; GDAL_CD39_NTPase.
DR Pfam; PF01150; GDAL_CD39; 1.
DR PROSITE; PS01238; GDAL_CD39_NTPASE; 1.
SQ SEQUENCE 467 AA; 51i64 MW; 8FCC200AA60D7376 CRC64;

Query Match 22.5%; Score 507.5; DB 10; Length 467;
Best Local Similarity 31.5%; Pred. No. 3.8e-35;
Matches 130; Conservative 72; Mismatches 162; Indels 49; Gaps 12;

QY 49 YGIMFDAGSGTCTRHVYTFVQKMPGQLPIL-----EGVEDSVKPGLSAFVDPQKQGAETV 104
Db 57 YAVIFDAGSGTSRVHVYRFNQ-----QLDLRIGHDLFLVTKPKLSAYAPENPEEAESL 112
QY 105 QGLLEAVKDSIPRSHWKKTPVVLKATAGLRLLPEHKAKALIFEVKEIFRKSPPFLVPKG-S 163
Db 113 VPLLEAEAVIPQELHPRTPVKVGATAGLRLEGDASNRILQAVSDMLKKRSTLKVEGDA 172
QY 164 VSIWTDQDEGIFAWTVNFTLQGLHGRQETVGTDLGGASTQITFLPQFEKTLKLE--QTP 221
Db 173 VSVLSGNQEGAYQWVTINYLNLGNLGHVSKTVAVVDLGGSGVQAYAISEDAAKAPQVP 232
QY 222 RG---YLTSPEMFNSTYKLYTHSVLGFGLKAARLATLGALETGDTGHTFRSACLPRWLE 278
Db 233 DGVESYITEMPLRGKKYLYVHSLYGLLAARAEVLKV-----SRDSENPCILSGFD 285
QY 279 AEWIFGGVKYQYGGNQGEGVGFPCYAEVLVRVR-----GKHLQHPVEVQGS 325
Db 286 GYTYGGVQYKATAPPSGS-SFSKQCNVLEALHVNATCSYKDCITFGIWNNGGGAGENN 344
QY 326 FYAFSYYYDRAVDTMID-YEKGILKVEFERKAREVCD-NLENFTS-----GSPF 375
Db 345 FFVASFFFEVDEAGFVDPNDANAIVRPVDPEDAACVACSTELKDLKSVFPRVKDGVY 404
QY 376 LCMDSLVTALLKDGFGFADSTVLQTKV-----NNLETGWALGATPHLQSL 424
Db 405*ICLDLVYQYTLVDVDFGDPQOEITLVRQIQYQDSLVEAAWPLGSATEAIESSL 457
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 09:33:34 ; Search time 38 Seconds  
(without alignments)

1787.762 Million cell updates/sec

Title: US-10-091-085-7

Perfect score: 2252

Sequence: 1 MATSWGTFFMLVVCVCSA.....ETGWLGAFTHLQLSLGISH 428

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_19Jun03.\*

- 1: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*
- 2: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*
- 3: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*
- 4: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*
- 5: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*
- 6: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*
- 7: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*
- 8: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*
- 9: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*
- 10: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*
- 11: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*
- 12: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*
- 13: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*
- 14: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*
- 15: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*
- 16: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*
- 17: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*
- 18: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*
- 19: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*
- 20: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*
- 21: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*
- 22: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*
- 23: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*
- 24: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2252	100.0	428	21	Human CD39-L4 vari
2	2252	100.0	428	22	Human CD39 like pr
3	2235	99.2	428	21	Human CD39-L4 prot
4	2235	99.2	428	22	Human CD39 like pr
5	2235	99.2	428	21	Human CD39-L4 vari
6	2235	99.2	428	22	Human CD39-L4 vari
7	2089	92.8	405	21	Human CD39-L66 pro
8	2089	92.8	405	22	Human CD39-L66 pro
9	1817.5	80.7	465	23	Mouse CD39L4 prote

10	1645	73.0	330	21	AAB53336	Human colon cancer
11	996	44.2	456	23	AAE19881	Human CD39L2 prote
12	996	44.2	484	22	AAE19881	Human CD39 like pr
13	993	44.1	463	23	ABB06124	Human NS protein s
14	993	44.1	467	23	ABJ04657	Protein of NOVX 15
15	989	43.9	456	22	AAE93329	Human polypeptide,
16	986	43.8	446	23	ABU04658	Protein of NOVX 15
17	894	39.7	462	22	AAU30882	Novel human secret
18	816.5	36.3	471	22	ABE66213	Mature human CD39
19	693.5	30.8	461	22	ABE66213	Drosophila melanog
20	693.5	30.8	464	22	ABE59611	Drosophila melanog
21	605	26.9	476	21	AAE70889	Protein encoded by
22	605	26.9	476	21	AAE70912	Human CD39-L4-2/3
23	548	24.3	476	21	AAE70888	Protein encoded by
24	548	24.3	476	21	AAE70911	Human CD39-L4-1 pr
25	503.5	22.4	467	20	AAW85687	DBX oligosaccharid
26	502.5	22.3	486	20	AAW85685	NBP46 root lectin.
27	502.5	22.3	486	23	AAU78819	Lotus japonicus le
28	493	21.9	462	20	AAW85684	NBP46 root lectin.
29	493	21.9	462	23	AAU78818	Dolichos biflorus
30	486.5	21.6	496	23	AAU78820	Medicago sativa le
31	484	21.5	455	22	AAE81952	Pea blight resista
32	483	21.4	472	24	ABP81286	Arabidopsis thalia
33	477.5	21.2	496	20	AAW85686	NBP46 root lectin.
34	380	16.9	495	20	AAE33296	Human membrane spa
35	371.5	16.5	458	22	AAU29271	Human PRO polypept
36	371.5	16.5	458	24	ABU71359	Human PRO5779 prot
37	371.5	16.5	458	24	ABU65816	Human secreted/tra
38	371.5	16.5	458	24	ABU66149	Novel human secret
39	371.5	16.5	458	24	ABU67653	Human secreted/tra
40	371.5	16.5	458	24	ABU65511	Human PRO polypept
41	371.5	16.5	458	24	ABU58647	Human PRO polypept
42	371.5	16.5	458	24	ABU56183	Human secreted/tra
43	371.5	16.5	458	24	ABU57178	Human PRO polypept
44	371.5	16.5	458	24	ABU10757	Human secreted/tra
45	369.5	16.4	474	21	AAE70900	Protein encoded by

## ALIGNMENTS

RESULT 1  
AA44850  
ID AAY44850 standard; Protein; 428 AA.

XX AC AAY44850;

XX DT 18-MAY-2000 (first entry)

XX Human CD39-L4 variant-ACR III mutant protein.

XX CD39-L4; human; apyrase; nucleotide diphosphatase; NDPase; variant;  
XX ATP diphosphohydrolase; ATPase; adenosine diphosphate; ADP; treatment;  
XX platelet aggregation; antithrombotic; thrombosis; myocardial infarction;  
XX cerebral ischaemia; angina; vascular graft; extracorporeal circulation;  
XX molecular weight marker; nutritional supplement; tumour; prevention;  
XX drug targeting; substitution mutation.

OS Homo sapiens.

OS Synthetic.

XX Key Location/Qualifiers

FT Misc-difference 168 /note= "Wild type Asp substituted with Thr"

FT Misc-difference 170 /note= "Wild type Ser substituted with Gln"

FT Misc-difference 175 /note= "Wild type Leu substituted with Phe"

XX WO200004041-A2.

XX 27-JAN-2000.

PD

XX

PF 16-JUL-1999; 99WO-US16180.  
 XX 16-JUL-1998; 98US-0118205.  
 PR 24-JUL-1998; 98US-0122449.  
 PR 04-FEB-1999; 99US-0244444.  
 PR 19-MAR-1999; 99US-0273447.  
 PR 09-JUL-1999; 99US-0350836.  
 XX (HYSE-) HYSEQ INC.  
 XX Ford J, Mulero J;  
 PI WPI; 2000-182397/16.  
 DR N-PSDB; AAZ50357.  
 XX New nucleic acid encoding human CD39-like protein, useful for treating  
 PT and preventing thrombotic disease -  
 XX Claim 17; Fig 6; 125pp; English.  
 PS The present amino acid sequence is the CD39-L4 variant, designated  
 XX as ACR III mutant protein, an apyrase and/or nucleotide diphosphatase  
 CC (NDPase). It is isolated from the human foetal liver-spleen cDNA library,  
 CC b2HFS20W. It is a soluble ATP diphosphohydrolase (ATPase) and is  
 CC involved in the hydrolysis of adenosine diphosphate (ADP), the agonist  
 CC that causes platelet aggregation. CD39-L4 protein has 30% and 80%  
 CC homology to human and murine CD39. It has platelet aggregation inhibition  
 CC and antithrombotic activity. CD39-L4 is used to treat or prevent  
 CC thrombosis, myocardial infarction, cerebral ischaemia and angina. It is  
 CC also used in vitro, to maintain vascular grafts or during extracorporeal  
 CC circulation, to hydrolyse NDP, as molecular weight markers and as  
 CC nutritional supplements. It is used to identify therapeutic agents that  
 CC bind and modulate CD39-L4. It is coupled to toxins for targeting drugs  
 CC to tumours or other cells that express CD39-L4.  
 XX Sequence 428 AA;  
 SQ Query Match 100.0%; Score 2252; DB 21; Length 428;  
 Best Local Similarity 100.0%; Pred. No. 7.4e-209;  
 Matches 428; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 MATSWGTVFVFWLVVSCVSAVSHRNQOTWPEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60  
 Db 1 MATSWGTVFVFWLVVSCVSAVSHRNQOTWPEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60  
 Qy 61 RIHVYTFVQKPGQLPILGEVDSVKGSLSAFVDQPKQGAETVQGLLEAKDSIPRSHW 120  
 Db 61 RIHVYTFVQKPGQLPILGEVDSVKGSLSAFVDQPKQGAETVQGLLEAKDSIPRSHW 120  
 Qy 121 KKTVPVLKATAGLRLLEHAKALLFEVKGIFRKSPLVPGKSVSINTGDEGIFAWVTV 180  
 Db 121 KKTVPVLKATAGLRLLEHAKALLFEVKGIFRKSPLVPGKSVSINTGDEGIFAWVTV 180  
 Qy 181 NFLTQLHGRHQETVGTLDLGGASTQITFLPQPEKTLQTPRGYLTSEFNFSTYKLYTH 240  
 Db 181 NFLTQLHGRHQETVGTLDLGGASTQITFLPQPEKTLQTPRGYLTSEFNFSTYKLYTH 240  
 Qy 241 SYLFGFLKAARLATLGALETGTDGHTFRSACLPRWLEAWIFGGVKYQYGGNQEYGVF 300  
 Db 241 SYLFGFLKAARLATLGALETGTDGHTFRSACLPRWLEAWIFGGVKYQYGGNQEYGVF 300  
 Qy 301 EPCVAEVLVRVGRKLHOPVEEVQSGSFYAFSYVDVADVTMDIDYKGGILKVEDFERKAR 360  
 Db 301 EPCVAEVLVRVGRKLHOPVEEVQSGSFYAFSYVDVADVTMDIDYKGGILKVEDFERKAR 360  
 Qy 361 EVCNLENFTSGSPFLCWDLSYITALLKDGFGFADSTVLQITKKVNNIETGALGATFHL 420  
 Db 361 EVCNLENFTSGSPFLCWDLSYITALLKDGFGFADSTVLQITKKVNNIETGALGATFHL 420  
 Qy 421 LQSLGISH 428  
 Db 421 LQSLGISH 428

RESULT 2  
 AAB72240  
 ID AAB72240 standard; Protein; 428 AA.  
 XX AAB72240;  
 AC AAB72240;  
 XX 14-MAY-2001 (first entry)  
 DT Human CD39 like protein CD39-L4 variant ACRIII amino acid sequence.  
 DE Human CD39-like protein; apyrase; NDPase; platelet function inhibitor;  
 XX myocardial infarction; cerebral ischaemia; angina; arterial thrombosis;  
 KW cerebral artery thrombosis; platelet aggregation; inflammation;  
 KW apoptosis; autoimmune disorder; neurological disorder; mutant; muten;  
 KW Alzheimer's disease; Parkinson's disease; cancer; CD39-L4.  
 XX Homo sapiens.  
 OS WO200110205-A1.  
 XX 15-FEB-2001.  
 PD 09-AUG-2000; 2000WO-US21790.  
 XX 09-AUG-1999; 99US-0370265.  
 PR 11-JAN-2000; 2000US-0481238.  
 PR 25-APR-2000; 2000US-0557800.  
 PR 28-MAY-2000; 2000US-0583231.  
 PR 30-JUN-2000; 2000US-0608285.  
 XX (HYSE-) HYSEQ INC.  
 XX Ford J, Mulero JJ, Yeung G;  
 PI WPI; 2001-147489/15.  
 DR N-PSDB; AAF63385.  
 XX Polynucleotides encoding human CD39-like polypeptides, with apyrase  
 PT and/or NDPase activity, which are useful in the treatment of  
 PT pathological conditions caused by thrombosis (e.g. myocardial  
 PT infarction) and inflammatory disorders -  
 XX Claim 19; Fig 6; 203pp; English.  
 CC This invention relates to polynucleotides encoding human CD39-like  
 CC polypeptides with apyrase and/or NDPase activity. The polypeptides having  
 CC ATPase, including NDPase, activity are useful for inhibiting platelet  
 CC function and can therefore be used in the prophylaxis or treatment of  
 CC pathological conditions caused by or involving thrombosis or excessive  
 CC coagulation or excessive platelet aggregation, such as myocardial  
 CC infarction, cerebral ischaemia, angina, arterial thrombosis, cerebral  
 CC artery thrombosis or intracardiac thrombosis, and conditions associated  
 CC with venous thrombosis. CD39-L4 and CD39-L2 polypeptides are useful in  
 CC modulating disease states (including platelet aggregation, inflammation  
 CC and apoptosis) associated with ADP or other purinergic signalling by  
 CC reducing the levels of NDPs. The polypeptides are also useful for  
 CC prophylaxis or treatment of inflammation related disorders, such as  
 CC disorders involving sepsis or systemic inflammatory response syndrome or  
 CC SIRS (and associated conditions such as fever, tachycardia, tachypnea,  
 CC cytokine overstimulation); autoimmune disorders such as thrombosis,  
 CC atherosclerosis, acute pancreatitis, dermatitis, including psoriasis,  
 CC cirrhosis, reperfusion injury, asthma, multiple sclerosis, arthritis;  
 CC neurological disorders including neurodegenerative diseases, epilepsy,  
 CC depression, Alzheimer's disease, Parkinson's disease, Huntington's  
 CC disease, and amyotrophic lateral sclerosis; and cancer. The present  
 CC sequence represents human CD39 like protein CD39-L4 variant ACRIII.  
 XX Sequence 428 AA;  
 SQ Query Match 100.0%; Score 2252; DB 22; Length 428;  
 Best Local Similarity 100.0%; Pred. No. 7.4e-209;  
 Matches 428; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MATSWGTFFMLVSVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT	60
DB	1	MATSWGTFFMLVSVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT	60
QY	61	RIHVYTFVKMPGQLPILGEVFDVSKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW	120
DB	61	RIHVYTFVKMPGQLPILGEVFDVSKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW	120
QY	121	KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPLVPKGSVIMTQDEGIFAWVTV	180
DB	121	KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPLVPKGSVIMTQDEGIFAWVTV	180
QY	181	NFLTQGLHGHROETVGTDLGGASTQITFLPQPEKTLQETPRGYLTSFEMFNSTYKLYTH	240
DB	181	NFLTQGLHGHROETVGTDLGGASTQITFLPQPEKTLQETPRGYLTSFEMFNSTYKLYTH	240
QY	241	SYLGFGLKAARLATLGALETGTGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEVEVGF	300
DB	241	SYLGFGLKAARLATLGALETGTGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEVEVGF	300
QY	301	EPCYAEVLVRVGRKGLHQPEEVQSGSFYASYVDRAVDTMDIDYEKGGILKVEDFERKAR	360
DB	301	EPCYAEVLVRVGRKGLHQPEEVQSGSFYASYVDRAVDTMDIDYEKGGILKVEDFERKAR	360
QY	361	EVCNDLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGVALGATFHL	420
DB	361	EVCNDLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGVALGATFHL	420
QY	421	LQSLGISH 428	
DB	421	LQSLGISH 428	
RESULT 3			
AA444849			
ID	AA444849	standard; Protein; 428 AA.	
XX	AC	AA444849;	
DT	18-MAY-2000	(first entry)	
XX	DE	Human CD39-L4 protein.	
XX	DE	CD39-L4; human; apyrase; nucleotide diphosphatase; NDPase;	
KW	KW	ATP diphosphohydrolase; ATPase; adenosine diphosphate; ADP; treatment;	
KW	KW	platelet aggregation; antithrombotic; thrombosis; myocardial infarction;	
KW	KW	cerebral ischaemia; angina; vascular graft; extracorporeal circulation;	
KW	KW	molecular weight marker; nutritional supplement; tumour; prevention;	
XX	XX	drug targeting; Apyrase Conserved Region; ACR.	
OS	OS	Homo sapiens.	
XX	XX	Key	Location/Qualifiers
FT	FT	Peptide	1..22
FT	FT	Protein	23..428
FT	FT	Protein	/label= Leader_peptide
FT	FT	Protein	/label= Mature_human_CD39-L4_protein
FT	FT	Protein	/note= "Homologous to human and murine CD39"
FT	FT	Binding-site	54..58
FT	FT	Region	/label= ATP_Binding_region
FT	FT	Region	129..134
FT	FT	Region	/label= Apyrase_Conserved_Region
FT	FT	Region	169..173
FT	FT	Region	/label= Apyrase_Conserved_Region
FT	FT	Region	199..206
FT	FT	Region	/note= "Conserved motif in ATPases"
XX	XX	WO200004041-A2.	
PN	PN	27-JAN-2000.	
PD	PD	16-JUL-1999;	99WO-US16180.
XX	XX		

XX	16-JUL-1998;	98US-0118205.	
PR	24-JUL-1998;	98US-0122449.	
PR	04-FEB-1999;	99US-0244444.	
PR	19-MAR-1999;	99US-0273447.	
PR	09-JUL-1999;	99US-0350836.	
XX	(HYSE-) HYSEQ INC.		
PA	Ford J, Mulero J;		
XX	WPI: 2000-182397/16.		
XX	N-PSDB; AA250356, AA250359.		
XX	New nucleic acid encoding human CD39-like protein, useful for treating		
PT	and preventing thrombotic disease -		
XX	Claim 15; Fig 2; 125pp; English.		
XX	The present amino acid sequence is the CD39-L4 protein, an apyrase		
CC	and/or nucleotide diphosphatase (NDPase). It is isolated from the human		
CC	foetal liver-glycine CDNA library, b2HPLS20W. It is a soluble ATP		
CC	diphosphohydrolase (ATPase) and is involved in the hydrolysis of		
CC	adenosine diphosphate (ADP), the agonist that causes platelet		
CC	aggregation. CD39-L4 protein has 30% and 80% homology to human and		
CC	murine CD39. It has platelet aggregation inhibition and antithrombotic		
CC	activity. CD39-L4 is used to treat or prevent thrombosis, myocardial		
CC	infarction, cerebral ischaemia and angina. It is also used in vitro, to		
CC	maintain vascular grafts or during extracorporeal circulation, to		
CC	hydrolyse NDP, as molecular weight markers and as nutritional		
CC	supplements. It is used to identify therapeutic agents that bind and		
CC	modulate CD39-L4. It is coupled to toxins for targeting drugs to tumours		
CC	or other cells that express CD39-L4.		
XX	Sequence 428 AA;		
SQ	Query Match	99.2%; Score 2235; DB 21; Length 428;	
	Best Local Similarity	99.3%; Pred. No. 3.3e-207;	
	Matches 425; Conservative	0; Mismatches 3; Indels 0; Gaps 0;	
QY	1	MATSWGTFFMLVSVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT	60
DB	1	MATSWGTFFMLVSVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT	60
QY	61	RIHVYTFVKMPGQLPILGEVFDVSKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW	120
DB	61	RIHVYTFVKMPGQLPILGEVFDVSKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW	120
QY	121	KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPLVPKGSVIMTQDEGIFAWVTV	180
DB	121	KKTPVVLKATAGLRLLPEHKAKALLFEVKEIFRKSPLVPKGSVIMTQDEGIFAWVTV	180
QY	181	NFLTQGLHGHROETVGTDLGGASTQITFLPQPEKTLQETPRGYLTSFEMFNSTYKLYTH	240
DB	181	NFLTQGLHGHROETVGTDLGGASTQITFLPQPEKTLQETPRGYLTSFEMFNSTYKLYTH	240
QY	241	SYLGFGLKAARLATLGALETGTGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEVEVGF	300
DB	241	SYLGFGLKAARLATLGALETGTGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEVEVGF	300
QY	301	EPCYAEVLVRVGRKGLHQPEEVQSGSFYASYVDRAVDTMDIDYEKGGILKVEDFERKAR	360
DB	301	EPCYAEVLVRVGRKGLHQPEEVQSGSFYASYVDRAVDTMDIDYEKGGILKVEDFERKAR	360
QY	361	EVCNDLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGVALGATFHL	420
DB	361	EVCNDLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTKKVNNIETGVALGATFHL	420
QY	421	LQSLGISH 428	
DB	421	LQSLGISH 428	

```

RESULT 4
AAB72238
ID AAB72238 standard; Protein; 428 AA.
XX
AC AAB72238;
XX
DT 14-MAY-2001 (first entry)
XX
DE Human CD39 like protein CD39-L4 amino acid sequence.
XX
KW Human CD39-like protein; apyrase; NDPase; platelet function inhibitor;
KW myocardial infarction; cerebral ischaemia; angina; arterial thrombosis;
KW cerebral artery thrombosis; platelet aggregation; inflammation;
KW apoptosis; autoimmune disorder; neurological disorder;
KW Alzheimer's disease; Parkinson's disease; cancer; CD39-L4.
XX
OS Homo sapiens.
XX
FN WO200110205-A1.
XX
PD 15-FEB-2001.
XX
PF 09-AUG-2000; 2000WO-US21790.
XX
PR 09-AUG-1999; 99US-0370265.
PR 11-JAN-2000; 2000US-0481238.
PR 25-APR-2000; 2000US-0557800.
PR 26-MAY-2000; 2000US-0583231.
PR 30-JUN-2000; 2000US-0608285.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Ford J, Mulero JJ, Yeung G;
XX
WPI; 2001-147489/15.
DR N-PSDB; AAF63383.
XX
Polynucleotides encoding human CD39-like polypeptides, with apyrase
and/or NDPase activity, which are useful in the treatment of
PT pathological conditions caused by thrombosis (e.g. myocardial
PT infarction) and inflammatory disorders -
XX
Claim 15; Fig 2; 203pp; English.
XX
PS This invention relates to polynucleotides encoding human CD39-like
CC polypeptides with apyrase and/or NDPase activity. The polypeptides having
CC ATPDase, including NDPase, activity are useful for inhibiting platelet
CC function and can therefore be used in the prophylaxis or treatment of
CC pathological conditions caused by or involving thrombosis or excessive
CC coagulation or excessive platelet aggregation, such as myocardial
CC infarction, cerebral ischaemia, angina, arterial thrombosis, cerebral
CC artery thrombosis or intracardiac thrombosis, and conditions associated
CC with venous thrombosis. CD39-L4 and CD39-L2 polypeptides are useful in
CC modulating disease states (including platelet aggregation, inflammation
CC and apoptosis) associated with ADP or other purinergic signalling by
CC reducing the levels of NDPs. The polypeptides are also useful for
CC prophylaxis or treatment of inflammation related disorders, such as
CC disorders involving sepsis or systemic inflammatory response syndrome or
CC SIRS (and associated conditions such as fever, tachycardia, tachypnea,
CC cytokine overstimulation); autoimmune disorders such as thrombosis,
CC atherosclerosis, acute pancreatitis, dermatitis, including psoriasis,
CC cirrhosis, reperfusion injury, asthma, multiple sclerosis, arthritis;
CC neurological disorders including neurodegenerative diseases, epilepsy,
CC depression, Alzheimer's disease, Parkinson's disease, Huntington's
CC disease, and amyotrophic lateral sclerosis; and cancer. The present
CC sequence represents human CD39 like protein CD39-L4.
XX
SQ Sequence 428 AA;
XX
Query Match 99.2%; Score 2235; DB 22; Length 428;
Best Local Similarity 99.3%; Pred. No. 3.3e-207;
Matches 425; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```



PT infarction) and inflammatory disorders -  
XX Example 9; Page 142-144; 203pp; English.

XX This invention relates to polynucleotides encoding human CD39-like  
XX polypeptides with apyrase and/or NTPase activity. The polypeptides having  
XX function and can therefore be used in the prophylaxis or treatment of  
XX pathological conditions caused by or involving thrombosis or excessive  
XX coagulation or excessive platelet aggregation, such as myocardial  
XX infarction, cerebral ischaemia, angina, arterial thrombosis, cerebral  
XX artery thrombosis or intracardiac thrombosis, and conditions associated  
XX with venous thrombosis. CD39-L4 and CD39-L2 polypeptides are useful in  
XX modulating disease states (including platelet aggregation, inflammation  
XX and apoptosis) associated with ADP or other purinergic signalling by  
XX reducing the levels of NTPs. The polypeptides are also useful for  
XX prophylaxis or treatment of inflammation related disorders, such as  
XX disorders involving sepsis or systemic inflammatory response syndrome or  
XX SIRS (and associated conditions such as fever, tachycardia, tachypnea,  
XX cytokine overstimulation); autoimmune disorders such as thrombosis,  
XX atherosclerosis, acute pancreatitis, dermatitis, including psoriasis,  
XX cirrhosis, reperfusion injury, asthma, multiple sclerosis, arthritis;  
XX neurological disorders including neurodegenerative diseases, epilepsy,  
XX depression, Alzheimer's disease, Parkinson's disease, Huntington's  
XX disease, and amyotrophic lateral sclerosis; and cancer. The present  
XX sequence represents the CD39 like protein CD39-L4 amino acid sequence.

SQ Sequence 428 AA;

Query Match 99.2%; Score 2235; DB 22; Length 428;  
Best Local Similarity 99.3%; Pred. No. 3.3e-207;  
Matches 425; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MATSWGTVFVFLVVCVCSAVSHRNQOTWFEGLFLSSMCPIVNSASTLYGIMFDAGSTGT 60  
DB 1 MATSWGTVFVFLVVCVCSAVSHRNQOTWFEGLFLSSMCPIVNSASTLYGIMFDAGSTGT 60  
QY 61 RIHVYTFVQKMPQLPILEGVDFSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120  
DB 61 RIHVYTFVQKMPQLPILEGVDFSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120  
QY 121 KTPVVLKATAGLURLPEHAKALLFEVKIIPKSPFLVPKGSVIMTQDEGIFAWVTV 180  
DB 121 KTPVVLKATAGLURLPEHAKALLFEVKIIPKSPFLVPKGSVIMTQDEGIFAWVTV 180  
QY 181 NLTGQLHGRQETVGTDLGGASTQITLPQPEKLTLEQTPRGVLTSEFMFNSTYKLYTH 240  
DB 181 NLTGQLHGRQETVGTDLGGASTQITLPQPEKLTLEQTPRGVLTSEFMFNSTYKLYTH 240  
QY 241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEVEGV 300  
DB 241 SYLGFGLKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEVEGV 300  
QY 301 EPCVAEVLVRVVRKGLHQPEVQVGSFYAFSYVDRAVDTMDIDYKGGILKVEDFERKAR 360  
DB 301 EPCVAEVLVRVVRKGLHQPEVQVGSFYAFSYVDRAVDTMDIDYKGGILKVEDFERKAR 360  
QY 361 EVCDNLENFTSGSPFLCMLDLYTALLKDGFGADSTVLQLTKKVNNIETGVALGATFHL 420  
DB 361 EVCDNLENFTSGSPFLCMLDLYTALLKDGFGADSTVLQLTKKVNNIETGVALGATFHL 420  
QY 421 LQSLGISH 428  
DB 421 LQSLGISH 428

RESULT 6  
AAE19883  
ID AAE19883 standard; Protein; 428 AA.  
XX AAE19883;  
XX 18-JUN-2002 (first entry)

XX DE Human CD39L4 protein.  
XX KW Human; CD39-like protein; CD39L4 protein; therapy; immune deficiency;  
KW autoimmune disorder; multiple sclerosis; systemic lupus erythematosus;  
KW rheumatoid arthritis; autoimmune thyroiditis; allergic reaction; asthma;  
KW insulin dependent diabetes mellitus; periodontal disease; osteoporosis;  
KW osteoarthritis; wound healing; tissue repair; Alzheimer's disease; ulcer;  
KW Parkinson's disease; amyotrophic lateral sclerosis; Huntington's disease;  
KW nervous system disease; nerve injury; ischaemia-reperfusion injury;  
KW endotoxin lethality; arthritis; nephritis; inflammatory bowel disease;  
KW Crohn's disease; viricide; antibacterial; antifungal; neuroprotective;  
KW dermatological; immunosuppressive; vulnary; nocotropic; anticonvulsant;  
KW antiinflammatory; nephrotropic; gastrointestinal; vasotropic.  
OS Homo sapiens.  
XX US6350447-B1.  
XX 26-FEB-2002.  
XX 29-JAN-1999; 99US-0240639.  
XX 29-JAN-1999; 99US-0240639.  
XX (HYSE-) HYSEQ INC.  
XX Chadwick BP, Frischauf A;  
XX WPI: 2002-215262/27.  
XX N-PSDB; AADJ1695.  
XX An isolated polypeptide with phosphohydrolase activity, designated  
XX CD39L2, useful to identify other proteins with which binding occurs or  
XX identify inhibitors and for treatment of, e.g., Alzheimer's, multiple  
XX sclerosis and osteoporosis -  
XX Example; Fig 7; 101pp; English.  
XX The present invention relates to novel proteins with phosphohydrolase  
XX activity, designated CD-39-like (CD39L) proteins and polynucleotides  
XX encoding such proteins. CD39L proteins are useful to treat infectious  
XX diseases caused by viral, bacterial, fungal or other infection that may  
XX be treatable with CD39L. They are useful in the treatment of various  
XX immune deficiencies and disorders, autoimmune disorders such as multiple  
XX sclerosis, systemic lupus erythematosus, rheumatoid arthritis, autoimmune  
XX thyroiditis and insulin dependent diabetes mellitus, allergic reactions  
XX and conditions such as asthma and other respiratory problems, periodontal  
XX disease, osteoporosis, osteoarthritis and other tooth repair processes.  
XX They may have utility in compositions used for bone, cartilage, tendon,  
XX ligament and/or nerve tissue growth or regeneration as well as for wound  
XX healing and tissue repair and replacement and in the treatment of burns,  
XX incisions and ulcers. CD39L proteins may also be useful for proliferation  
XX of neural cells and for regeneration of nerve and brain tissue, i.e. for  
XX the treatment of central nervous system diseases such as Alzheimer's  
XX disease, Parkinson's disease, amyotrophic lateral sclerosis, Huntington's  
XX disease, peripheral nervous system diseases peripheral nerve injuries,  
XX peripheral neuropathy and localised neuropathies. They are also used to  
XX treat mechanical and traumatic disorders which involve degeneration,  
XX death or trauma to neural cells or nerve tissue. CD39L proteins of non-  
XX invention are also useful to promote better or faster closure of non-  
XX healing wounds, including pressure ulcers, ulcers associated with  
XX vascular insufficiency and surgical and traumatic wounds. They also  
XX exhibit anti-inflammatory activity and may be used to treat inflammatory  
XX conditions including chronic or acute conditions), including ischaemia-  
XX reperfusion injury, endotoxin lethality, arthritis, nephritis, cytokine  
XX or chemokine-induced lung injury, inflammatory bowel disease or Crohn's  
XX disease. The present sequence is human CD39L4 protein.  
XX Sequence 428 AA;

Query Match 99.2%; Score 2235; DB 23; Length 428;  
Best Local Similarity 99.3%; Pred. No. 3.3e-207;

Matches 425; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MATSWGTVFVFLVSVCSAVSHRNQQTWFGIFLSSMCPINVSASTLYGIMFDAGSTGT 60  
DB 1 MATSWGTVFVFLVSVCSAVSHRNQQTWFGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

QY 61 RIHYVTVFQKMPGQLPILGEVFDVSRKGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120  
DB 61 RIHYVTVFQKMPGQLPILGEVFDVSRKGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120

QY 121 KKTVPVLKATAGLRLLEHAKALLFEVKEIFRKSPLVPKGSVINTGDEGIFAWTV 180  
DB 121 KKTVPVLKATAGLRLLEHAKALLFEVKEIFRKSPLVPKGSVINTGDEGIFAWTV 180

QY 181 NFLTGQHGHRQETVGTDLGGASTQITFLPQPEKTLQTPRGYLTSGFEMFNSTYKLYTH 240  
DB 181 NFLTGQHGHRQETVGTDLGGASTQITFLPQPEKTLQTPRGYLTSGFEMFNSTYKLYTH 240

QY 241 SYLGFGGLKAARLATLGALETGDTGHTFRSACLPRWLEAEWIFGVKYQYGGNQEVEVF 300  
DB 241 SYLGFGGLKAARLATLGALETGDTGHTFRSACLPRWLEAEWIFGVKYQYGGNQEVEVF 300

QY 301 EPCYAEVLVRVVRGKLHQPVEVQSGSYAFSYDDRAVDTDMDIYKGGILKVEDFERKAR 360  
DB 301 EPCYAEVLVRVVRGKLHQPVEVQSGSYAFSYDDRAVDTDMDIYKGGILKVEDFERKAR 360

QY 361 EVCNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTCKVNNIETGWLATFHL 420  
DB 361 EVCNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTCKVNNIETGWLATFHL 420

QY 421 LQSLGISH 428  
DB 421 LQSLGISH 428

RESULT 7  
AAY44851  
ID AAY44851 standard; Protein; 405 AA.

XX AC AAY44851;  
XX DE  
DT 18-MAY-2000 (first entry)  
XX DE  
XX KW Human CD39-L66 protein.  
XX KW CD39-L66; human; CD39-L66; apyrase; nucleotide diphosphatase; NDPase;  
XX KW ATP Diphosphohydrolase; ATPase; adenosine diphosphate; ADP; treatment;  
XX KW platelet aggregation; antithrombotic; thrombosis; myocardial infarction;  
XX KW cerebral ischaemia; angina; vascular graft; extracorporeal circulation;  
XX KW molecular weight marker; nutritional supplement; tumour; prevention;  
XX KW drug targeting; splice variant.

XX OS Homo sapiens.  
XX PN WO200004041-A2.  
XX PD 27-JAN-2000.  
XX PF 16-JUL-1999; 99WO-US16180.  
XX PR 16-JUL-1998; 98US-0118205.  
XX PR 24-JUL-1998; 98US-0122449.  
XX PR 04-FEB-1999; 99US-024444.  
XX PR 19-MAR-1999; 99US-0273447.  
XX PR 09-JUL-1999; 99US-0350836.  
XX PA (HYSE-) HYSEQ INC.  
XX PI Ford J, Mulero J;  
XX PI WPI; 2000-182397/16.  
XX DR N-FSDS; AAZ50358.  
XX XX

PT New nucleic acid encoding human CD39-like protein, useful for treating  
XX and preventing thrombotic disease -  
PS Claim 15; Page 124-125; 125pp; English.  
XX

CC The present amino acid sequence is the CD39-L66 protein, a splice  
CC variant of the CD39-L4 protein. It is an apyrase and/or nucleotide  
CC diphosphatase (NDPase), isolated from the human foetal liver-spleen cDNA  
CC library, b2HFLS20W. It is a soluble ATP Diphosphohydrolases (ATPDase) and  
CC is involved in the hydrolysis of adenosine diphosphate (ADP), the agonist  
CC that causes platelet aggregation. CD39-L4 protein has 30% and 80%  
CC homology to human and murine CD39. It has platelet aggregation inhibition  
CC and antithrombotic activity. CD39-L4 is used to treat or prevent  
CC thrombosis, myocardial infarction, cerebral ischaemia and angina. It is  
CC also used, in vitro, to maintain vascular grafts or during extracorporeal  
CC circulation, to hydrolyse NDP, as molecular weight markers and as  
CC nutritional supplements. It is used to identify therapeutic agents that  
CC bind and modulate CD39-L4. It is coupled to toxins for targeting drugs  
CC to tumours or other cells that express CD39-L4.

SQ Sequence 405 AA;

Query Match 92.8%; Score 2089; DB 21; Length 405;  
Best Local Similarity 99.2%; Pred. No. 4e-193;  
Matches 397; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MATSWGTVFVFLVSVCSAVSHRNQQTWFGIFLSSMCPINVSASTLYGIMFDAGSTGT 60  
DB 1 MATSWGTVFVFLVSVCSAVSHRNQQTWFGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

QY 61 RIHYVTVFQKMPGQLPILGEVFDVSRKGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120  
DB 61 RIHYVTVFQKMPGQLPILGEVFDVSRKGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120

QY 121 KKTVPVLKATAGLRLLEHAKALLFEVKEIFRKSPLVPKGSVINTGDEGIFAWTV 180  
DB 121 KKTVPVLKATAGLRLLEHAKALLFEVKEIFRKSPLVPKGSVINTGDEGIFAWTV 180

QY 181 NFLTGQHGHRQETVGTDLGGASTQITFLPQPEKTLQTPRGYLTSGFEMFNSTYKLYTH 240  
DB 181 NFLTGQHGHRQETVGTDLGGASTQITFLPQPEKTLQTPRGYLTSGFEMFNSTYKLYTH 240

QY 241 SYLGFGGLKAARLATLGALETGDTGHTFRSACLPRWLEAEWIFGVKYQYGGNQEVEVF 300  
DB 241 SYLGFGGLKAARLATLGALETGDTGHTFRSACLPRWLEAEWIFGVKYQYGGNQEVEVF 300

QY 301 EPCYAEVLVRVVRGKLHQPVEVQSGSYAFSYDDRAVDTDMDIYKGGILKVEDFERKAR 360  
DB 301 EPCYAEVLVRVVRGKLHQPVEVQSGSYAFSYDDRAVDTDMDIYKGGILKVEDFERKAR 360

QY 361 EVCNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQ 400  
DB 361 EVCNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQ 400

RESULT 8  
AAB72239  
ID AAB72239 standard; Protein; 405 AA.  
XX AC AAB72239;  
XX DT 14-MAY-2001 (first entry)  
XX DE Human CD39 like protein CD39-L4 amino acid sequence.  
XX KW Human CD39-like protein; apyrase; NDPase; platelet function inhibitor;  
XX KW myocardial infarction; cerebral ischaemia; angina; arterial thrombosis;  
XX KW cerebral artery thrombosis; platelet aggregation; inflammation;  
XX KW apoptosis; autoimmune disorder; neurological disorder;  
XX KW Alzheimer's disease; Parkinson's disease; cancer; CD39-L4.  
XX OS Homo sapiens.  
XX XX





DT 18-JUN-2002 (first entry)  
DE Human CD39L2 protein.  
XX Human; CD-39-like protein; CD39L2 protein; immune deficiency;  
XX autoimmune disorder; multiple sclerosis; systemic lupus erythematosus;  
XX rheumatoid arthritis; autoimmune thyroiditis; allergic reaction; asthma;  
XX insulin dependent diabetes mellitus; periodontal disease; osteoporosis;  
XX osteoarthritis; wound healing; tissue repair; Alzheimer's disease; ulcer;  
XX Parkinson's disease; amyotrophic lateral sclerosis; Huntington's disease;  
XX nervous system disease; nerve injury; ischaemia-reperfusion injury;  
XX endotoxin lethality; arthritis; nephritis; inflammatory bowel disease;  
XX Crohn's disease; virucide; antibacterial; antifungal; neuroprotective;  
XX dermatological; immunosuppressive; vulnerary; nootropic; anticonvulsant;  
XX antiinflammatory; nephrotropic; gastrointestinal; vasotropic.  
OS Homo sapiens.  
XX US6350447-B1.  
XX 26-FEB-2002.  
XX 29-JAN-1999; 99US-0240639.  
XX 29-JAN-1999; 99US-0240639.  
XX (HYSE-) HYSEQ INC.  
XX Chadwick BP, Frischauf A;  
XX WPI; 2002-215262/27.  
XX N-PSDB; AAD31693.  
XX An isolated polypeptide with phosphohydrolase activity, designated  
XX CD39L2, useful to identify other proteins with which binding occurs or  
XX identify inhibitors and for treatment of, e.g., Alzheimer's, multiple  
XX sclerosis and osteoporosis -  
XX Claim 1; Fig 4; 101pp; English.  
XX The present invention relates to novel proteins with phosphohydrolase  
XX activity, designated CD-39-like (CD39L) proteins and polynucleotides  
XX encoding such proteins. CD39L proteins are useful to treat infectious  
XX diseases caused by viral, bacterial, fungal or other infection that may  
XX be treatable with CD39L. They are useful in the treatment of various  
XX immune deficiencies and disorders, autoimmune disorders such as multiple  
XX sclerosis, systemic lupus erythematosus, rheumatoid arthritis, autoimmune  
XX thyroiditis and insulin dependent diabetes mellitus, allergic reactions  
XX and conditions such as asthma and other respiratory problems, periodontal  
XX disease, osteoporosis, osteoarthritis and other tooth repair processes.  
XX They may have utility in compositions used for bone, cartilage, tendon,  
XX ligament and/or nerve tissue growth or regeneration as well as for wound  
XX healing and tissue repair and replacement and in the treatment of burns,  
XX incisions and ulcers. CD39L proteins may also be useful for proliferation  
XX of neural cells and for regeneration of nerve and brain tissue, i.e. for  
XX the treatment of central nervous system diseases such as Alzheimer's  
XX disease, Parkinson's disease, amyotrophic lateral sclerosis, Huntington's  
XX disease, peripheral nervous system diseases peripheral nerve injuries,  
XX treat mechanical and traumatic disorders which involve degeneration,  
XX death or trauma to neural cells or nerve tissue. CD39L proteins of the  
XX invention are also useful to promote better or faster closure of non-  
XX healing wounds, including pressure ulcers, ulcers associated with  
XX vascular insufficiency and surgical and traumatic wounds. They also  
XX exhibit anti-inflammatory activity and may be used to treat inflammatory  
XX conditions including chronic or acute conditions), including ischaemia-  
XX reperfusion injury, endotoxin lethality, arthritis, nephritis, cytokine  
XX or chemokine-induced lung injury, inflammatory bowel disease or Crohn's  
XX disease. The present sequence is human CD39L2 protein.  
SQ Sequence 456 AA;  
Query Match 44.2%; Score 996; DB 23; Length 456;

Best Local Similarity 52.4%; Pred. No. 2.6e-87;  
Matches 204; Conservative 54; Mismatches 125; Indels 6; Gaps 4;  
QY 40 PINVSA---STLYGIMFDAGSTGTRIHYVTFVQKMPGQLPILGEVFDSDVKPGLSAFVDQ 96  
DB 62 PLGTAADGHEVFGIMFDAGSTGTRVHVQFT-RPPRETPTLTHTFKAVKPGLSAVADD 120  
QY 97 PKGASTVOGLLEVAKDSIPRSHWKTTPVVLKATAGLRLLPEHAKALLFEVKEIFRKPSP 156  
DB 121 VEKSAQIGIRELLDVAQDIPDFWFKATPLVLKATAGLRLLPEHAKALLQKVKVEVPKASP 180  
QY 157 FLVPKGSVIMTGQDSGIFAWTVNFLTGLGHGROETVGTLDLGGASTQITFLPOPEKT 216  
DB 181 FLVGGDCVIMMGTDGVSAMITINLTGSLKTPGGSSVGMLDLGGSTQIAFLPRVEGT 240  
QY 217 LEOTPRGYLTSPFMENSTYKLYTHSYLGFGLKAARLATLGALETE-GTDGHTFRSACLPR 275  
DB 241 LQASPPGYLTALRMFNRTYKLYSYLGLGLMSARLAILGGVGPQAKDGKELVSPCLSP 300  
QY 276 WLEAEWIFGGVYQYCGNQEVEGPEPCYAEVLRVVRGKHLQDPEEVQVRSFYAFSYYDR 335  
DB 301 SPKGEWEHAEVTVRVSGQKAAASLHELCAARVSEVLQNRVHRTEEVKHVDVYAFSYYDL 360  
QY 336 AVDTDMIDYKGGILKVEDFERKAREVCNLENFTSGSPFLCNDLSVITALLKDGFGFAD 395  
DB 361 AAGVGLIDAEKGGSLVVGDFEIAAKYVCRVLTETQPOSSPSCMDLTVLSLLQ8-FGPPR 419  
QY 396 STVLQLTKKVNNIETGALCATPHLLQSL 424  
DB 420 SKVLKTRKIDNVETSWALGAIFHYIDSL 448  
RESULT 12  
AAB72241  
ID AAB72241 standard; Protein; 484 AA.  
XX AAB72241;  
XX 14-MAY-2001 (first entry)  
XX Human CD39 like protein CD39-L2 amino acid sequence.  
XX Human CD39-like protein; apyrase; NDPase; platelet function inhibitor;  
XX myocardial infarction; cerebral ischaemia; angina; arterial thrombosis;  
XX cerebral artery thrombosis; platelet aggregation; inflammation;  
XX apoptosis; autoimmune disorder; neurological disorder;  
XX Alzheimer's disease; Parkinson's disease; cancer; CD39-L2.  
XX Homo sapiens.  
XX WO200110205-A1.  
XX 15-FEB-2001.  
XX 09-AUG-2000; 2000WO-US21790.  
XX 09-AUG-1999; 99US-0370265.  
XX 11-JAN-2000; 2000US-0481238.  
XX 25-APR-2000; 2000US-0557800.  
XX 26-MAY-2000; 2000US-0583231.  
XX 30-JUN-2000; 2000US-0608285.  
XX (HYSE-) HYSEQ INC.  
XX Ford J, Mulero JJ, Yeung G;  
XX WPI; 2001-147489/15.  
XX N-PSDB; AAF63386.  
XX Polynucleotides encoding human CD39-like polypeptides, with apyrase  
XX and/or NDPase activity, which are useful in the treatment of  
XX pathological conditions caused by thrombosis (e.g. myocardial  
XX infarction) and inflammatory disorders -

XX Claim 39; Page 162-164; 203pp; English.  
 XX This invention relates to polynucleotides encoding human CD39-like  
 CC polypeptides with apyrase and/or NTPase activity. The polypeptides have  
 CC ATPase, including NTPase, activity are useful for inhibiting platelet  
 CC function and can therefore be used in the prophylaxis or treatment of  
 CC pathological conditions caused by or involving thrombosis or excessive  
 CC coagulation or excessive platelet aggregation, such as myocardial  
 CC infarction, cerebral ischaemia, angina, arterial thrombosis, cerebral  
 CC artery thrombosis or intracardiac thrombosis, and conditions associated  
 CC with venous thrombosis. CD39-L4 and CD39-L2 polypeptides are useful in  
 CC modulating disease states (including platelet aggregation, inflammation  
 CC and apoptosis) associated with ADP or other purinergic signalling by  
 CC reducing the levels of NTPs. The polypeptides are also useful for  
 CC prophylaxis or treatment of inflammation related disorders, such as  
 CC disorders involving sepsis or systemic inflammatory response syndrome or  
 CC SIRS (and associated conditions such as fever, tachycardia, tachypnea,  
 CC cytokine overstimulation); autoimmune disorders such as thrombosis,  
 CC atherosclerosis, acute pancreatitis, dermatitis, including psoriasis,  
 CC cirrhosis, reperfusion injury, asthma, multiple sclerosis, arthritis;  
 CC neurological disorders including neurodegenerative diseases, epilepsy,  
 CC depression, Alzheimer's disease, Parkinson's disease, Huntington's  
 CC disease, and amyotrophic lateral sclerosis; and cancer. The present  
 CC sequence represents human CD39 like protein CD39-L2.  
 XX  
 SQ Sequence 484 AA;

Query Match 44.2%; Score 996; DB 22; Length 484;  
 Best Local Similarity 52.4%; Pred. No. 2.8e-87;  
 Matches 204; Conservative 54; Mismatches 125; Indels 6; Gaps 4;

QY 40 PINVSA---STLYGIMFDAGSTGTRHVTYFVQKMGQPLILEGEVFDSPKGLSAFVDQ 96  
 DB 90 PLGTAADGHEVYFYGIMFDAGSTGTRHVFQRT-RPPRETPTLTHTFKAVKPGLSAYDD 148  
 QY 97 PKQAEVTVQGLLEVAKDSIPRSHWKTTPVVLKATAGLRLPEHKAALLFEVKEIFRKPSP 156  
 DB 149 VEKSAQGIREDLDAKQDIPDFWKATPLVLKATAGLRLPEHKAALLFEVKEIFRKPSP 208  
 QY 157 FLVPGKGSVIMTQDEGIFAWVTNFLTQGLHGHROETVGTLDLGGASTQITFLPQPEKT 216  
 DB 209 FLVGDGCVSIMGNGTDEGSAWITNFLTQGLHGHROETVGTLDLGGASTQITFLPQPEKT 268  
 QY 217 LEOTPRGVLTFEFNFSTYKLYTHSYLGFGLKAARLALTLGALETE-GTGDHTFSACLPR 275  
 DB 269 LQASPPGVLTRALNFNRYTKLYSYLGLGLMSARLALTLGVGEQPKAGKELVSPCLSP 328  
 QY 276 WLEAEWIFGGVYQYQGNQGEVGFPCYAEVLVRVVRGKLHQPEVQSGFSAFYFYDR 335  
 DB 329 SFKGEWEHAEVTVRVSGQKAAASLHELCAARVSEVLQNRVHRTVEVKHVDVFAFSYYDL 388  
 QY 336 AVDTMDIDYEGGILKVEDFERKAREVCDNLENTSGSPFLCMLDSITALLKDGFGFAD 395  
 DB 389 AAGVGLIDAEKGGSLVGDGFPAIAKYVCTLETQPSQSPFCMDLTVVSLLLQD-FGFPFR 447  
 QY 396 STVLQLTQKVNITGWTGALGATPHLQSL 424  
 DB 448 SKVLTKRKIDNVETSWALGAIHFYIDSL 476

RESULT 13  
 ID ABB06124 standard; Protein; 463 AA.  
 XX ABB06124  
 AC ABB06124;  
 XX  
 DT 10-MAY-2002 (first entry)  
 XX Human NS protein sequence SEQ ID NO:216.  
 XX Human; cytostatic; osteopathic; gynaecological; neuroprotective;  
 KW antirheumatic; antiarthritic; antipsoriatic; ophthalmological; anti-HIV;

KW vasotropic; antiarteriosclerotic; antiinflammatory; dermatological;  
 KW anorectic; muscular; antinfertility; cardiovascular; anticoagulant;  
 KW antifibrinolytic; hypotension; antiasthmatic; immunomodulator; cardiant;  
 KW anticonvulsant; antidiabetic; tranquilliser; antidepressant; neuroleptic;  
 KW gastrointestinal; virucide; antiulcer; cancer; osteoporosis; dystonia;  
 KW contraceptive; vaccine; gene therapy; cancer; osteoporosis; dystonia;  
 KW endometriosis; degenerative disease; multiple sclerosis; psoriasis;  
 KW rheumatoid arthritis; cataract; restenosis; atherosclerosis; glaucoma;  
 KW inflammation; skin disorder; obesity; muscular dystrophy; AIDS;  
 KW infertility; cardiovascular disease; coagulation disease; hypertension;  
 KW ischaemia; asthma; immune disease; epilepsy; angina; neurodegeneration;  
 KW diabetic; anxiety; depression; schizophrenia; viral disease; stroke;  
 KW gastric ulcer; Alzheimer's disease.

OS Homo sapiens.  
 XX W0200206315-A2.  
 XX 24-JAN-2002.  
 XX 17-JUL-2001; 2001WO-IL00653.  
 XX 18-JUL-2000; 2000IL-0137345.  
 XX 13-DEC-2000; 2000IL-0140354.  
 XX (COMP-) COMPUGEN LTD.

PI Mintz L, Freilich S, Bernstein J;  
 XX WPI; 2002-155037/20.  
 DR N-PSDB; ABL39778.  
 XX One hundred and twenty eight novel nucleic acid sequences, useful for  
 PT treating and diagnosing e.g. cancer, asthma and Alzheimer's -  
 XX  
 PS Claim 6; Page 251-253; 290pp; English.

XX ABL39691 to ABL39818 represent novel human nucleic acid sequences  
 CC encoding the proteins given in ABB06037 to ABB06164. The novel sequences  
 CC (NS) can have cytostatic, osteopathic, gynaecological, neuroprotective,  
 CC antirheumatic, antiarthritic, antipsoriatic, ophthalmological, virucide,  
 CC vasotropic, antiarteriosclerotic, antiinflammatory, dermatological,  
 CC anorectic, muscular, anti-HIV, antinfertility, cardiovascular,  
 CC anticoagulant, antifibrinolytic, hypotension, antiasthmatic, cardiant,  
 CC immunomodulator, anticonvulsant, antidiabetic, tranquilliser, antiulcer,  
 CC antidepressant, gastrointestinal, neuroleptic, cerebroprotective,  
 CC nootropic and contraceptive activities. The NS can be used in vaccines, and  
 CC gene therapy and antiseize therapy. Nucleic acids, expression vectors and  
 CC antibodies from the present invention can be used for treating and  
 CC diagnosing e.g. cancer, osteoporosis, endometriosis, degenerative  
 CC diseases, dystonia, multiple sclerosis, rheumatoid arthritis, psoriasis,  
 CC cataracts, restenosis, atherosclerosis, inflammation, skin disorders,  
 CC glaucoma, obesity, muscular dystrophy, AIDS, infertility, cardiovascular  
 CC disease, coagulation disease, ischaemia, hypertension, asthma, immune  
 CC disease, epilepsy, angina, neurodegeneration, diabetes, anxiety,  
 CC depression, schizophrenia, viral disease, gastric ulcers, stroke,  
 CC Alzheimer's disease and as a contraceptive.

SQ Sequence 463 AA;

Query Match 44.1%; Score 993; DB 23; Length 463;  
 Best Local Similarity 52.2%; Pred. No. 5.2e-87;  
 Matches 203; Conservative 55; Mismatches 125; Indels 6; Gaps 4;

QY 40 PINVSA---STLYGIMFDAGSTGTRHVTYFVQKMGQPLILEGEVFDSPKGLSAFVDQ 96  
 DB 69 PLGTAADGHEVYFYGIMFDAGSTGTRHVFQRT-RPPRETPTLTHTFKAVKPGLSAYDD 127  
 QY 97 PKQAEVTVQGLLEVAKDSIPRSHWKTTPVVLKATAGLRLPEHKAALLFEVKEIFRKPSP 156  
 DB 128 VEKSAQGIREDLDAKQDIPDFWKATPLVLKATAGLRLPEHKAALLFEVKEIFRKPSP 187  
 QY 157 FLVPGKGSVIMTQDEGIFAWVTNFLTQGLHGHROETVGTLDLGGASTQITFLPQPEKT 216

DB 188 FLVGDDCVSIMNGTDEGVSAMITINFLTGSLSKTPGSSVGMGLDLGGSTQIAFLPRVEGT 247  
QY 217 LEQTPRGYLTSFEMFNSTYKLYTHSYLGFGLKAARLATLGALETE-GTGDHTFRSACLPR 275  
DB 248 LQASPPGYLTALRMFNRTYKLYSYLGLGMSARLAILGGVSGQPAKDGKELVSPCLSP 307  
QY 276 WLEAEWIFGGVKYQYCGNQEVEGFPFCYAEVLVRVGRKLHQPEEVORGSGFYAFSYYDR 335  
DB 308 SPKGEWEHAEVTVRVSGQAAASLHELCAARVSEVLQNRVHRTEEVKVDVDFYAFSYYDL 367  
QY 336 AVDTDMIDYKGGILKVEDPERKAREVCNLENFTSGSPFLCMLDLSYITALLKDGFGFAD 395  
DB 368 AAGVGLIDAEKGGSLVVGDFEIAAKYVCRTELETQPOSSPFCMDLTVSLLQES-FGPPR 426  
QY 396 STVLQTLTKKVNNTETGWCALGATPHLLQSL 424  
DB 427 SKVLKTRKIDNVETSWALGAIFHYIDSL 455

RESULT 14  
ABJ04657  
ID ABJ04657 standard; Protein; 467 AA.  
XX AC ABJ04657;  
XX DT 11-OCT-2002 (first entry)  
XX DE Protein of NOVX 15a SEQ ID NO 36.  
XX KW Cytostatic; antidiabetic; anorectic; metabolic; nootropic; antilipaemic;  
KW neuroprotective; antiparkinsonian; anticonvulsant; cerebroprotective;  
KW tranquilliser; neuroleptic; antidiabetic; antiulcer; antiinflammatory;  
KW anti-HIV; antiallergic; antirheumatic; antiarthritic; NOVX; diabetes;  
KW metabolic disorder; obesity; infectious disease; Alzheimer's disease;  
KW anorexia; neurodegenerative disorder; Parkinson's disorder; obesity;  
KW immune disorder; haematopoietic disorder; dyslipidaemia; chronic disease;  
KW metabolic syndrome X; wasting disorder; cancer; neurological disorder;  
KW epilepsy; stroke; mental disorder; schizophrenia; goiter;  
KW vesicular transport; cystic fibrosis; gastrointestinal disorder;  
KW diabetes mellitus; ulcerative colitis; AIDS; allergic reaction;  
KW multiple sclerosis; rheumatoid arthritis; AIDS; transgenic animal;  
KW gene therapy.  
XX OS Unidentified.  
XX PN WO200246409-A2.  
XX PD 13-JUN-2002.  
XX PF 06-DEC-2001; 2001WO-US46586.  
XX PR 06-DEC-2000; 2000US-251660P.  
PR 12-DEC-2000; 2000US-255029P.  
PR 08-JAN-2001; 2001US-260326P.  
PR 24-JAN-2001; 2001US-263800P.  
PR 20-FEB-2001; 2001US-269942P.  
PR 24-APR-2001; 2001US-286183P.  
PR 20-AUG-2001; 2001US-313627P.  
PR 12-SEP-2001; 2001US-318712P.  
XX (CURA-) CURAGEN CORP.  
XX PA Guo X, Li L, Paturajan M, Shimkets RA, Casman SJ, Malyankar UM;  
PI Tchernev VT, Vernet CAM, Spytek KA, Shenoy SG, Alsobrook JP;  
PI Edinger S, Peyman JA, Stone DJ, Ellerman K, Gangolli EA;  
PI Boldog FL, Colman SD, Eisen AJ, Liu X, Padigaru M, Spaderna SK;  
PI Zerhusen BD;  
XX WP1; 2002-547774/58.  
DR N-PSDB; ABT05470.  
XX Novel isolated polypeptide, designated NOVX, useful for treating or

PT preventing cancer, diabetes, obesity, dyslipidemia, anorexia, and  
PT metabolic, neurodegenerative, immune and hematopoietic disorders -  
XX Claim 1; Page 140; 421pp; English.  
XX The invention relates to an isolated polypeptide, designated NOVX,  
CC comprising a sequence fully defined in the specification. The isolated  
CC protein, its encoding polynucleotide or an antibody created from the  
CC syndrome is useful in the manufacture of a medicament for treating a  
CC syndrome associated with a human disease, preferably a NOVX-associated  
CC disorder, or for treating or preventing a NOVX-associated disorder in a  
CC subject, preferably human. The isolated protein, its encoding  
CC polynucleotide or an antibody created from the protein are also useful  
CC for treating or preventing metabolic disorders, diabetes, obesity,  
CC infectious disease, anorexia, neurodegenerative disorder, Alzheimer's  
CC disease, Parkinson's disorder, immune disorders, haematopoietic  
CC disorders, and various dyslipidaemias, metabolic disturbances associated  
CC with obesity, the metabolic syndrome X, wasting disorders associated with  
CC chronic diseases, and cancer. The isolated protein, its encoding  
CC polynucleotide or an antibody created from the protein are useful for  
CC treating or preventing neurological disorders such as epilepsy, stroke,  
CC mental disorders including mood, anxiety, schizophrenic disorders,  
CC disorders of vesicular transport such as cystic fibrosis, diabetes  
CC mellitus, goiter, gastrointestinal disorders including ulcerative  
CC colitis, other conditions associated with abnormal vesicle trafficking  
CC including AIDS, allergic reactions, multiple sclerosis and rheumatoid  
CC arthritis. A cell comprising the vector of the invention is useful for  
CC producing non-human transgenic animals. The polynucleotide of the  
CC invention can be used to treat disorders by gene therapy. This sequence  
CC represents one of the isolated NOVX proteins of the invention.  
XX SQ Sequence 467 AA;  
Query Match 44.1%; Score 993; DB 23; Length 467;  
Best Local Similarity 52.2%; Pred. No. 5.2e-87;  
Matches 203; Conservative 55; Mismatches 125; Indels 6; Gaps 4;  
QY 40 PINVSA---STLYGIMFDAGSTGTRIHVYTFVQKMPGQLPILGEVFDVSKPLSAPVDQ 96  
DB 73 PLGTAADGHEVFYGMFDAGSTGTRVHVQFT-RPPRETPTLTHTFKALKPLSAYADD 131  
QY 97 PKGASTVOGLLEVAKDSIPRSHWKTTPVLKATAGLRLLPEHKAKALLPEVKEIFRKSP 156  
DB 132 VEKSAQGIREDLDVAKQDIPDFWKATPLVLKATAGLRLLPGEKAQKLQKKEVFKASP 191  
QY 157 FLVPKGSVSIIMGQDEGIPAWTVNFLTQOLHGHROETVGTLDLGGASTQITFLPOPEKT 216  
DB 192 FLVGDDCVSIMNGTDEGVSAMITINFLTGSLSKTPGSSVGMGLDLGGSTQIAFLPRVEGT 251  
QY 217 LEQTPRGYLTSFEMFNSTYKLYTHSYLGFGLKAARLATLGALETE-GTGDHTFRSACLPR 275  
DB 252 LQASPPGYLTALRMFNRTYKLYSYLGLGMSARLAILGGVSGQPAKDGKELVSPCLSP 311  
QY 276 WLEAEWIFGGVKYQYCGNQEVEGFPFCYAEVLVRVGRKLHQPEEVORGSGFYAFSYYDR 335  
DB 312 SPKGEWEHAEVTVRVSGQAAASLHELCAARVSEVLQNRVHRTEEVKVDVDFYAFSYYDL 371  
QY 336 AVDTDMIDYKGGILKVEDPERKAREVCNLENFTSGSPFLCMLDLSYITALLKDGFGFAD 395  
DB 372 AAGVGLIDAEKGGSLVVGDFEIAAKYVCRTELETQPOSSPFCMDLTVSLLQES-FGPPR 430  
QY 396 STVLQTLTKKVNNTETGWCALGATPHLLQSL 424  
DB 431 SKVLKTRKIDNVETSWALGAIFHYIDSL 459  
RESULT 15  
AAM93929  
ID AAM93929 standard; Protein; 456 AA.  
XX AC AAM93929;  
XX DT 06-NOV-2001 (first entry)





GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 10:02:15 ; Search time 15 Seconds  
(without alignments)  
1207.270 Million cell updates/sec

Title: US-10-091-085-7  
Perfect score: 2252  
Sequence: 1 MATSWGTVFPMVLVSCVCSA.....ETGWLGAFTHLQLSLGISH 428  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5  
Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA: \*  
1: /cgn2\_6/ptodata/1/iaa/5A COMB.pdp.\*  
2: /cgn2\_6/ptodata/1/iaa/5B COMB.pdp.\*  
3: /cgn2\_6/ptodata/1/iaa/6A COMB.pdp.\*  
4: /cgn2\_6/ptodata/1/iaa/6B COMB.pdp.\*  
5: /cgn2\_6/ptodata/1/iaa/PTUS COMB.pdp.\*  
6: /cgn2\_6/ptodata/1/iaa/backfiles.pdp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2252	100.0	428	4	US-09-608-285A-7
2	2252	100.0	428	4	US-09-350-836B-7
3	2252	100.0	428	4	US-09-370-265-7
4	2252	100.0	428	4	US-09-557-800C-7
5	2235	99.2	428	4	US-09-608-285A-3
6	2235	99.2	428	4	US-09-608-285A-5
7	2235	99.2	428	4	US-09-240-639-6
8	2235	99.2	428	4	US-09-240-639-9
9	2235	99.2	428	4	US-09-350-836B-3
10	2235	99.2	428	4	US-09-350-836B-5
11	2235	99.2	428	4	US-09-370-265-3
12	2235	99.2	428	4	US-09-370-265-5
13	2235	99.2	428	4	US-09-557-800C-3
14	2235	99.2	428	4	US-09-557-800C-5
15	2089	92.8	405	4	US-09-608-285A-25
16	2089	92.8	405	4	US-09-370-265-25
17	2089	92.8	405	4	US-09-557-800C-25
18	1822.5	80.9	465	4	US-09-557-800C-56
19	1817.5	80.7	465	4	US-09-240-639-8
20	996	44.2	456	4	US-09-240-639-2
21	996	44.2	484	4	US-09-608-285A-27
22	996	44.2	484	4	US-09-370-265-27
23	996	44.2	484	4	US-09-557-800C-27
24	816.5	36.3	471	4	US-09-608-285A-60
25	507.5	22.5	467	4	US-09-129-112-19
26	502.5	22.3	459	4	US-09-129-112-9
27	493	21.9	462	4	US-09-129-112-2

28	486.5	21.6	462	4	US-09-129-112-15	Sequence 15, Appl
29	484	21.5	455	4	US-09-240-639-10	Sequence 10, Appl
30	472.5	21.0	473	4	US-09-240-639-12	Sequence 12, Appl
31	441.5	19.6	454	4	US-09-240-639-11	Sequence 11, Appl
32	366	16.3	502	4	US-09-557-800C-55	Sequence 55, Appl
33	366	16.3	510	3	US-08-930-921-1	Sequence 1, Appl
34	357.5	15.9	529	4	US-09-240-639-4	Sequence 4, Appl
35	229	10.2	148	4	US-09-240-639-17	Sequence 17, Appl
36	201	8.9	153	4	US-09-240-639-13	Sequence 13, Appl
37	191	8.5	153	4	US-09-240-639-15	Sequence 15, Appl
38	189.5	8.4	150	4	US-09-240-639-16	Sequence 16, Appl
39	189	8.4	154	4	US-09-240-639-14	Sequence 14, Appl
40	106.5	4.7	339	4	US-09-252-991A-20854	Sequence 20854, A
41	104.5	4.6	553	1	US-08-565-386-12	Sequence 12, Appl
42	103.5	4.6	535	3	US-09-137-077-2	Sequence 2, Appl
43	95.5	4.2	510	4	US-09-252-991A-17138	Sequence 17138, A
44	94	4.2	506	4	US-09-134-001C-4383	Sequence 4383, Ap
45	91.5	4.1	334	4	US-09-252-991A-22395	Sequence 22395, A

ALIGNMENTS

RESULT 1  
US-09-608-285A-7  
; Sequence 7, Application US/09608285A  
; Patent No. 6335013  
; GENERAL INFORMATION:  
; APPLICANT: Ford, John  
; APPLICANT: Mulero, Julio  
; APPLICANT: Yeung, George  
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO CD39-LIKE  
; FILE OF INVENTION: POLYPEPTIDES  
; FILE REFERENCE: 28110/36570  
; CURRENT APPLICATION NUMBER: US/09/608, 285A  
; CURRENT FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: 09/583, 231  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 09/557, 800  
; PRIOR FILING DATE: 2000-04-25  
; PRIOR APPLICATION NUMBER: 09/481, 238  
; PRIOR FILING DATE: 2000-01-11  
; PRIOR APPLICATION NUMBER: 09/370, 265  
; PRIOR FILING DATE: 1999-08-09  
; PRIOR APPLICATION NUMBER: PCT/US99/16180  
; PRIOR FILING DATE: 1999-07-16  
; PRIOR APPLICATION NUMBER: 09/350, 836  
; PRIOR FILING DATE: 1999-07-09  
; PRIOR APPLICATION NUMBER: 09/273, 447  
; PRIOR FILING DATE: 1999-03-19  
; PRIOR APPLICATION NUMBER: 09/244, 444  
; PRIOR FILING DATE: 1999-02-04  
; PRIOR APPLICATION NUMBER: 09/122, 449  
; PRIOR FILING DATE: 1998-07-24  
; PRIOR APPLICATION NUMBER: 09/118, 205  
; PRIOR FILING DATE: 1998-07-16  
; NUMBER OF SEQ ID NOS: 60  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 7  
; LENGTH: 428  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-608-285A-7

Query Match	100.0%	Score	2252	DB	4	Length	428
Best Local Similarity	100.0%	Pred. No.	4.4e-247				
Matches	428	Conservative	0	Mismatches	0	Indels	0
Qy	1	MATSWGTVFPMVLVSCVCSAVSHRNQOTWFEGLFSMCPINVSASTLYGIMFDAGSTGT	60				
Db	1	MATSWGTVFPMVLVSCVCSAVSHRNQOTWFEGLFSMCPINVSASTLYGIMFDAGSTGT	60				
Qy	61	RIHVTYFQKMPQLFILEGEVDSVKPGLSAFVDQPKOGAETVQGLLEVAKDSIPRSHW	120				

*method highly significant*



Db 241 SYLGFLKAARLATLGALETEGDTGHTFRSACLPRWLEAEWIFGKVYQYGGNQEVEGF 300  
Qy 301 EPCVAEVLVRGKLHQPEVQSGSFYAFSYYYDRAVDTDMDIDYKGGILKVEDFERKAR 360  
Db 301 EPCVAEVLVRGKLHQPEVQSGSFYAFSYYYDRAVDTDMDIDYKGGILKVEDFERKAR 360  
Qy 361 EVCNLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQLTKKVNNIETGMAFGATFHL 420  
Db 361 EVCNLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQLTKKVNNIETGMAFGATFHL 420  
Qy 421 LOSLGISH 428  
Db 421 LOSLGISH 428

RESULT 4

US-09-557-800C-7  
; Sequence 7, Application US/09557800C  
; Patent No. 6476211  
; GENERAL INFORMATION:  
; APPLICANT: Ford, John  
; APPLICANT: Mulero, Julio  
; APPLICANT: Yeung, George  
; TITLE OF INVENTION: Methods and Materials Relating to CD39-Like  
; TITLE OF INVENTION: Polypeptides  
; FILE REFERENCE: 28110/36457  
; CURRENT APPLICATION NUMBER: US/09/557,800C  
; CURRENT FILING DATE: 2000-04-25  
; PRIOR APPLICATION NUMBER: 09/481,238  
; PRIOR FILING DATE: 2000-01-11  
; PRIOR APPLICATION NUMBER: 09/370,265  
; PRIOR FILING DATE: 1999-08-09  
; PRIOR APPLICATION NUMBER: PCT/US99/16180  
; PRIOR FILING DATE: 1999-07-16  
; PRIOR APPLICATION NUMBER: 09/350836  
; PRIOR FILING DATE: 1999-07-09  
; PRIOR APPLICATION NUMBER: 09/273447  
; PRIOR FILING DATE: 1999-03-19  
; PRIOR APPLICATION NUMBER: 09/122449  
; PRIOR FILING DATE: 1998-07-24  
; PRIOR APPLICATION NUMBER: 09/244444  
; PRIOR FILING DATE: 1999-02-04  
; PRIOR APPLICATION NUMBER: 09/118,205  
; PRIOR FILING DATE: 1998-07-16  
; NUMBER OF SEQ ID NOS: 56  
; SOFTWARE: PatentIn ver. 2.0  
; SEQ ID NO 7  
; LENGTH: 428  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-557-800C-7

Query Match 100.0%; Score 2252; DB 4; Length 428;  
Best Local Similarity 100.0%; Pred. No. 4.4e-247;  
Matches 428; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 MATSWGTVFVFLVSVCSAVSHRNQOTWPEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60  
Db 1 MATSWGTVFVFLVSVCSAVSHRNQOTWPEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60  
Qy 61 RIHYVTVFVQKMPGQLPILGEVFDVSVPKGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120  
Db 61 RIHYVTVFVQKMPGQLPILGEVFDVSVPKGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120  
Qy 121 KKTVPVVKATAGLRLPEHAKALLFEVKEIFRKSPLVPKGSVSIIMTQDEGIFAWVTY 180  
Db 121 KKTVPVVKATAGLRLPEHAKALLFEVKEIFRKSPLVPKGSVSIIMTQDEGIFAWVTY 180  
Qy 181 NFLTQGLHGRQETVGTLDLGGASTQITFLPQPEKTLTQTPRGVLTSEFNFSTYKLYTH 240  
Db 181 NFLTQGLHGRQETVGTLDLGGASTQITFLPQPEKTLTQTPRGVLTSEFNFSTYKLYTH 240

Qy 241 SYLGFLKAARLATLGALETEGDTGHTFRSACLPRWLEAEWIFGKVYQYGGNQEVEGF 300  
Db 241 SYLGFLKAARLATLGALETEGDTGHTFRSACLPRWLEAEWIFGKVYQYGGNQEVEGF 300  
Qy 301 EPCVAEVLVRGKLHQPEVQSGSFYAFSYYYDRAVDTDMDIDYKGGILKVEDFERKAR 360  
Db 301 EPCVAEVLVRGKLHQPEVQSGSFYAFSYYYDRAVDTDMDIDYKGGILKVEDFERKAR 360  
Qy 361 EVCNLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQLTKKVNNIETGMAFGATFHL 420  
Db 361 EVCNLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQLTKKVNNIETGMAFGATFHL 420  
Qy 421 LOSLGISH 428  
Db 421 LOSLGISH 428

RESULT 5

US-09-608-285A-3  
; Sequence 3, Application US/09608285A  
; Patent No. 6335013  
; GENERAL INFORMATION:  
; APPLICANT: Ford, John  
; APPLICANT: Mulero, Julio  
; APPLICANT: Yeung, George  
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO CD39-LIKE  
; TITLE OF INVENTION: POLYPEPTIDES  
; FILE REFERENCE: 28110/36570  
; CURRENT APPLICATION NUMBER: US/09/608,285A  
; CURRENT FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: 09/583,231  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 09/557,800  
; PRIOR FILING DATE: 2000-04-25  
; PRIOR APPLICATION NUMBER: 09/481,238  
; PRIOR FILING DATE: 2000-01-11  
; PRIOR APPLICATION NUMBER: 09/370,265  
; PRIOR FILING DATE: 1999-08-09  
; PRIOR APPLICATION NUMBER: PCT/US99/16180  
; PRIOR FILING DATE: 1999-07-16  
; PRIOR APPLICATION NUMBER: 09/350,836  
; PRIOR FILING DATE: 1999-07-09  
; PRIOR APPLICATION NUMBER: 09/273,447  
; PRIOR FILING DATE: 1999-03-19  
; PRIOR APPLICATION NUMBER: 09/244,444  
; PRIOR FILING DATE: 1999-02-04  
; PRIOR APPLICATION NUMBER: 09/122,449  
; PRIOR FILING DATE: 1998-07-24  
; PRIOR APPLICATION NUMBER: 09/118,205  
; PRIOR FILING DATE: 1998-07-16  
; NUMBER OF SEQ ID NOS: 60  
; SOFTWARE: PatentIn ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 428  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-608-285A-3

Query Match 99.2%; Score 2235; DB 4; Length 428;  
Best Local Similarity 99.3%; Pred. No. 3.7e-245;  
Matches 425; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MATSWGTVFVFLVSVCSAVSHRNQOTWPEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60  
Db 1 MATSWGTVFVFLVSVCSAVSHRNQOTWPEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60  
Qy 61 RIHYVTVFVQKMPGQLPILGEVFDVSVPKGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120  
Db 61 RIHYVTVFVQKMPGQLPILGEVFDVSVPKGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120  
Qy 121 KKTVPVVKATAGLRLPEHAKALLFEVKEIFRKSPLVPKGSVSIIMTQDEGIFAWVTY 180  
Db 121 KKTVPVVKATAGLRLPEHAKALLFEVKEIFRKSPLVPKGSVSIIMTQDEGIFAWVTY 180

QY 181 NLTGQLHGRQETVGTLDLGASTQITFLPQFEKLEQTPRGYLTSFEMFNSTYKLYTH 240  
DB 181 NLTGQLHGRQETVGTLDLGASTQITFLPQFEKLEQTPRGYLTSFEMFNSTYKLYTH 240  
QY 241 SYLGFGKLAARLTLGALETGTGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEVEGF 300  
DB 241 SYLGFGKLAARLTLGALETGTGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEVEGF 300  
QY 301 EPCVAEVLVRVGRKGLHQPVEVQSGSFYAFSYDDRAVDTMDIDYKGGILKVEDFERKAR 360  
DB 301 EPCVAEVLVRVGRKGLHQPVEVQSGSFYAFSYDDRAVDTMDIDYKGGILKVEDFERKAR 360  
QY 361 EVCNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQTKKVNNIETGALGATFHL 420  
DB 361 EVCNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQTKKVNNIETGALGATFHL 420  
QY 421 LQSLGISH 428  
DB 421 LQSLGISH 428

## RESULT 6

US-09-608-285A-5  
; Sequence 5, Application US/09608285A  
; Patent No. 6335013

GENERAL INFORMATION:

APPLICANT: Ford, John

APPLICANT: Mulero, Julio

APPLICANT: Yeung, George

TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO CD39-LIKE

TITLE OF INVENTION: POLYPEPTIDES

FILE REFERENCE: 28110/36570

CURRENT APPLICATION NUMBER: US/09/608,285A

CURRENT FILING DATE: 2000-06-30

PRIOR APPLICATION NUMBER: 09/583,231

PRIOR FILING DATE: 2000-05-26

PRIOR APPLICATION NUMBER: 09/557,800

PRIOR FILING DATE: 2000-04-25

PRIOR APPLICATION NUMBER: 09/481,238

PRIOR FILING DATE: 2000-01-11

PRIOR APPLICATION NUMBER: 09/370,265

PRIOR FILING DATE: 1999-08-09

PRIOR APPLICATION NUMBER: PCT/US99/16180

PRIOR FILING DATE: 1999-07-16

PRIOR APPLICATION NUMBER: 09/350,836

PRIOR FILING DATE: 1999-07-09

PRIOR APPLICATION NUMBER: 09/273,447

PRIOR FILING DATE: 1999-03-19

PRIOR APPLICATION NUMBER: 09/244,444

PRIOR FILING DATE: 1999-02-04

PRIOR APPLICATION NUMBER: 09/122,449

PRIOR FILING DATE: 1998-07-24

PRIOR APPLICATION NUMBER: 09/118,205

PRIOR FILING DATE: 1998-07-16

NUMBER OF SEQ ID NOS: 60

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 5

LENGTH: 428

TYPE: PRT

ORGANISM: Homo sapiens

US-09-608-285A-5

Query Match 99.2%; Score 2235; DB 4; Length 428;  
Best Local Similarity 99.3%; Pred. No. 3.7e-245;  
Matches 425; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MATSWGTVFFMLVVCVCSAVSHRNQQTWPEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60  
DB 1 MATSWGTVFFMLVVCVCSAVSHRNQQTWPEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60  
QY 61 RIHYVTFVQKMPQQLPILEGEVFDVSKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120

DB 61 RIHYVTFVQKMPQQLPILEGEVFDVSKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120  
QY 121 KKTVPVVLKATAGLRLLEPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMTQODEGIFAWVTV 180  
DB 121 KKTVPVVLKATAGLRLLEPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMTQODEGIFAWVTV 180  
QY 181 NLTGQLHGRQETVGTLDLGASTQITFLPQFEKLEQTPRGYLTSFEMFNSTYKLYTH 240  
DB 181 NLTGQLHGRQETVGTLDLGASTQITFLPQFEKLEQTPRGYLTSFEMFNSTYKLYTH 240  
QY 241 SYLGFGKLAARLTLGALETGTGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEVEGF 300  
DB 241 SYLGFGKLAARLTLGALETGTGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEVEGF 300  
QY 301 EPCVAEVLVRVGRKGLHQPVEVQSGSFYAFSYDDRAVDTMDIDYKGGILKVEDFERKAR 360  
DB 301 EPCVAEVLVRVGRKGLHQPVEVQSGSFYAFSYDDRAVDTMDIDYKGGILKVEDFERKAR 360  
QY 361 EVCNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQTKKVNNIETGALGATFHL 420  
DB 361 EVCNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQTKKVNNIETGALGATFHL 420  
QY 421 LQSLGISH 428  
DB 421 LQSLGISH 428

## RESULT 7

US-09-240-639-6

; Sequence 6, Application US/09240639

; Patent No. 6350447

GENERAL INFORMATION:

APPLICANT: Chadwick, Brian Paul

APPLICANT: Frischauf, Anna-Maria

TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO CD39-LIKE

TITLE OF INVENTION: POLYPEPTIDES AND NUCLEIC ACIDS

FILE REFERENCE: 9598-066

CURRENT APPLICATION NUMBER: US/09/240,639

CURRENT FILING DATE: 1998-01-29

NUMBER OF SEQ ID NOS: 29

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 6

LENGTH: 428

TYPE: PRT

ORGANISM: Homo sapiens

US-09-240-639-6

Query Match 99.2%; Score 2235; DB 4; Length 428;  
Best Local Similarity 99.3%; Pred. No. 3.7e-245;  
Matches 425; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MATSWGTVFFMLVVCVCSAVSHRNQQTWPEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60  
DB 1 MATSWGTVFFMLVVCVCSAVSHRNQQTWPEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60  
QY 61 RIHYVTFVQKMPQQLPILEGEVFDVSKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120  
DB 61 RIHYVTFVQKMPQQLPILEGEVFDVSKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120  
QY 121 KKTVPVVLKATAGLRLLEPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMTQODEGIFAWVTV 180  
DB 121 KKTVPVVLKATAGLRLLEPEHKAKALLFEVKEIFRKSPFLVPKGSVSIMTQODEGIFAWVTV 180  
QY 181 NLTGQLHGRQETVGTLDLGASTQITFLPQFEKLEQTPRGYLTSFEMFNSTYKLYTH 240  
DB 181 NLTGQLHGRQETVGTLDLGASTQITFLPQFEKLEQTPRGYLTSFEMFNSTYKLYTH 240  
QY 241 SYLGFGKLAARLTLGALETGTGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEVEGF 300  
DB 241 SYLGFGKLAARLTLGALETGTGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEVEGF 300  
QY 301 EPCVAEVLVRVGRKGLHQPVEVQSGSFYAFSYDDRAVDTMDIDYKGGILKVEDFERKAR 360

Db 301 EPCYAEVLVRGKHLQHPVEVQSGSFYAFSYDDRAVDMDIDYKGGILKVEDFERKAR 360  
QY 361 EVCDNLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQLTKKVNNIETGVALGATFHL 420  
Db 361 EVCDNLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQLTKKVNNIETGVALGATFHL 420  
QY 421 LQSLGISH 428  
Db 421 LQSLGISH 428  
RESULT 8  
US-09-240-639-9  
; Sequence 9, Application US/09240639  
; Patent No. 6350447  
; GENERAL INFORMATION:  
; APPLICANT: Chadwick, Brian Paul  
; APPLICANT: Frischauf, Anna-Maria  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO CD39-LIKE  
; FILE OF INVENTION: POLYPEPTIDES AND NUCLEIC ACIDS  
; FILE REFERENCE: 9598-066  
; CURRENT APPLICATION NUMBER: US/09/240,639  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 9  
; LENGTH: 428  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-240-639-9

Query Match 99.2%; Score 2235; DB 4; Length 428;  
Best Local Similarity 99.3%; Pred. No. 3,7e-245;  
Matches 425; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 1 MATSWGTVPFVPMVLVSCVSAVSHRNQQTWPEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60  
Db 1 MATSWGTVPFVPMVLVSCVSAVSHRNQQTWPEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60  
QY 61 RIHYVTFVQKMPQLPILGEVFDVSKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120  
Db 61 RIHYVTFVQKMPQLPILGEVFDVSKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120  
QY 121 KKTVPVVKATAGLRLPEHKAKALLFEVKEIFRKSPLVPKGSVIMTQDEGIFAWTV 180  
Db 121 KKTVPVVKATAGLRLPEHKAKALLFEVKEIFRKSPLVPKGSVIMTQDEGIFAWTV 180  
QY 181 NFLTQGLHGRQETVGTDLGGASTQITFLPQPEKTLQTPRGYLTSPFENFNSTYKLYTH 240  
Db 181 NFLTQGLHGRQETVGTDLGGASTQITFLPQPEKTLQTPRGYLTSPFENFNSTYKLYTH 240  
QY 241 SYLGFGGLKAARLATLGALETGDTGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEVEVGF 300  
Db 241 SYLGFGGLKAARLATLGALETGDTGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEVEVGF 300  
QY 301 EPCYAEVLVRGKHLQHPVEVQSGSFYAFSYDDRAVDMDIDYKGGILKVEDFERKAR 360  
Db 301 EPCYAEVLVRGKHLQHPVEVQSGSFYAFSYDDRAVDMDIDYKGGILKVEDFERKAR 360  
QY 361 EVCDNLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQLTKKVNNIETGVALGATFHL 420  
Db 361 EVCDNLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQLTKKVNNIETGVALGATFHL 420  
QY 421 LQSLGISH 428  
Db 421 LQSLGISH 428

RESULT 9  
US-09-350-836B-3  
; Sequence 3, Application US/09350836B  
; Patent No. 6387645  
; GENERAL INFORMATION:  
; APPLICANT: Ford, John  
; APPLICANT: Mulero, Julio  
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE  
; FILE REFERENCE: 28110/35761  
; CURRENT APPLICATION NUMBER: US/09/350,836B  
; CURRENT FILING DATE: 1999-07-09  
; PRIOR APPLICATION NUMBER: 09/273,447

; APPLICANT: Ford, John  
; APPLICANT: Mulero, Julio  
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE  
; FILE OF INVENTION: POLYPEPTIDES  
; FILE REFERENCE: 28110/35761  
; CURRENT APPLICATION NUMBER: US/09/350,836B  
; CURRENT FILING DATE: 1999-07-09  
; PRIOR APPLICATION NUMBER: 09/273,447  
; PRIOR FILING DATE: 1999-03-19  
; PRIOR APPLICATION NUMBER: 09/118,205  
; PRIOR FILING DATE: 1998-07-16  
; PRIOR APPLICATION NUMBER: 09/122,449  
; PRIOR FILING DATE: 1998-07-24  
; PRIOR APPLICATION NUMBER: 09/244,444  
; PRIOR FILING DATE: 1999-02-04  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 428  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-350-836B-3  
Query Match 99.2%; Score 2235; DB 4; Length 428;  
Best Local Similarity 99.3%; Pred. No. 3,7e-245;  
Matches 425; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 1 MATSWGTVPFVPMVLVSCVSAVSHRNQQTWPEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60  
Db 1 MATSWGTVPFVPMVLVSCVSAVSHRNQQTWPEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60  
QY 61 RIHYVTFVQKMPQLPILGEVFDVSKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120  
Db 61 RIHYVTFVQKMPQLPILGEVFDVSKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120  
QY 121 KKTVPVVKATAGLRLPEHKAKALLFEVKEIFRKSPLVPKGSVIMTQDEGIFAWTV 180  
Db 121 KKTVPVVKATAGLRLPEHKAKALLFEVKEIFRKSPLVPKGSVIMTQDEGIFAWTV 180  
QY 181 NFLTQGLHGRQETVGTDLGGASTQITFLPQPEKTLQTPRGYLTSPFENFNSTYKLYTH 240  
Db 181 NFLTQGLHGRQETVGTDLGGASTQITFLPQPEKTLQTPRGYLTSPFENFNSTYKLYTH 240  
QY 241 SYLGFGGLKAARLATLGALETGDTGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEVEVGF 300  
Db 241 SYLGFGGLKAARLATLGALETGDTGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEVEVGF 300  
QY 301 EPCYAEVLVRGKHLQHPVEVQSGSFYAFSYDDRAVDMDIDYKGGILKVEDFERKAR 360  
Db 301 EPCYAEVLVRGKHLQHPVEVQSGSFYAFSYDDRAVDMDIDYKGGILKVEDFERKAR 360  
QY 361 EVCDNLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQLTKKVNNIETGVALGATFHL 420  
Db 361 EVCDNLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQLTKKVNNIETGVALGATFHL 420  
QY 421 LQSLGISH 428  
Db 421 LQSLGISH 428

RESULT 10  
US-09-350-836B-5  
; Sequence 5, Application US/09350836B  
; Patent No. 6387645  
; GENERAL INFORMATION:  
; APPLICANT: Ford, John  
; APPLICANT: Mulero, Julio  
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE  
; FILE REFERENCE: 28110/35761  
; CURRENT APPLICATION NUMBER: US/09/350,836B  
; CURRENT FILING DATE: 1999-07-09  
; PRIOR APPLICATION NUMBER: 09/273,447

```
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: 09/118,205
; PRIOR FILING DATE: 1998-07-16
; PRIOR APPLICATION NUMBER: 09/122,449
; PRIOR FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: 09/244,444
; PRIOR FILING DATE: 1999-02-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 5
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-350-836B-5

Query Match      99.2%; Score 2235; DB 4; Length 428;
Best Local Similarity 99.3%; Pred. No. 3.7e-245;
Matches 425; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MATSWGTVFVFLVVCVCSAVSHRNQQTWPEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
Db 1 MATSWGTVFVFLVVCVCSAVSHRNQQTWPEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

Qy 61 RIHYTTFVQKMPGQLPILEGEVFSVCKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
Db 61 RIHYTTFVQKMPGQLPILEGEVFSVCKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120

Qy 121 KKTPTVVLKATAGLRLLEPHKAKALLFEVKEIFRKSPPFLVPKGSVIMTGOGEI FAWTV 180
Db 121 KKTPTVVLKATAGLRLLEPHKAKALLFEVKEIFRKSPPFLVPKGSVIMTGOGEI FAWTV 180

Qy 181 NFLTQGLHGRHQRTVTGTLGGASTQITFLPQPEKTELEQTPRGVLTSEFNFNSTYKLYTH 240
Db 181 NFLTQGLHGRHQRTVTGTLGGASTQITFLPQPEKTELEQTPRGVLTSEFNFNSTYKLYTH 240

Qy 241 SYLGFGLKAARLATLGALETGTGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGBVGF 300
Db 241 SYLGFGLKAARLATLGALETGTGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGBVGF 300

Qy 301 EPCYAEVLVRVVRGKLHOPEEVQVRSFYAFSYYYDRAVDTMDIDYKGGILKVEDFERKAR 360
Db 301 EPCYAEVLVRVVRGKLHOPEEVQVRSFYAFSYYYDRAVDTMDIDYKGGILKVEDFERKAR 360

Qy 361 EVCNLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQLTKKVNNIETGVALGATFHL 420
Db 361 EVCNLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQLTKKVNNIETGVALGATFHL 420

Qy 421 LQSLGISH 428
Db 421 LQSLGISH 428

RESULT 11
US-09-370-265-3
; Sequence 3, Application US/09370265
; Patent No. 644771
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; APPLICANT: Mulero, Julio
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE
; FILE REFERENCE: 28111/35908
; CURRENT APPLICATION NUMBER: US/09/370,265
; CURRENT FILING DATE: 1999-08-09
; EARLIER APPLICATION NUMBER: PCT/US99/16180
; EARLIER FILING DATE: 1999-07-16
; EARLIER APPLICATION NUMBER: 09/350,836
; EARLIER FILING DATE: 1999-07-09
; EARLIER APPLICATION NUMBER: 09/273,447
; EARLIER FILING DATE: 1999-03-19
; EARLIER APPLICATION NUMBER: 09/350,836
; EARLIER FILING DATE: 1999-07-09
; EARLIER APPLICATION NUMBER: 09/244,444
; EARLIER FILING DATE: 1999-02-04
; EARLIER APPLICATION NUMBER: 09/273,447
; EARLIER FILING DATE: 1999-03-19
; EARLIER APPLICATION NUMBER: 09/244,444
; EARLIER FILING DATE: 1999-02-04
; EARLIER APPLICATION NUMBER: 09/122,449
; EARLIER FILING DATE: 1998-07-24
; EARLIER APPLICATION NUMBER: 09/118,205
; EARLIER FILING DATE: 1998-07-16
; NUMBER OF SEQ ID NOS: 37

; PRIOR FILING DATE: 1998-07-24
; EARLIER APPLICATION NUMBER: 09/118,205
; EARLIER FILING DATE: 1998-07-16
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 3
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-370-265-3

Query Match      99.2%; Score 2235; DB 4; Length 428;
Best Local Similarity 99.3%; Pred. No. 3.7e-245;
Matches 425; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 1 MATSWGTVFVFLVVCVCSAVSHRNQQTWPEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
Db 1 MATSWGTVFVFLVVCVCSAVSHRNQQTWPEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

Qy 61 RIHYTTFVQKMPGQLPILEGEVFSVCKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
Db 61 RIHYTTFVQKMPGQLPILEGEVFSVCKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120

Qy 121 KKTPTVVLKATAGLRLLEPHKAKALLFEVKEIFRKSPPFLVPKGSVIMTGOGEI FAWTV 180
Db 121 KKTPTVVLKATAGLRLLEPHKAKALLFEVKEIFRKSPPFLVPKGSVIMTGOGEI FAWTV 180

Qy 181 NFLTQGLHGRHQRTVTGTLGGASTQITFLPQPEKTELEQTPRGVLTSEFNFNSTYKLYTH 240
Db 181 NFLTQGLHGRHQRTVTGTLGGASTQITFLPQPEKTELEQTPRGVLTSEFNFNSTYKLYTH 240

Qy 241 SYLGFGLKAARLATLGALETGTGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGBVGF 300
Db 241 SYLGFGLKAARLATLGALETGTGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEGBVGF 300

Qy 301 EPCYAEVLVRVVRGKLHOPEEVQVRSFYAFSYYYDRAVDTMDIDYKGGILKVEDFERKAR 360
Db 301 EPCYAEVLVRVVRGKLHOPEEVQVRSFYAFSYYYDRAVDTMDIDYKGGILKVEDFERKAR 360

Qy 361 EVCNLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQLTKKVNNIETGVALGATFHL 420
Db 361 EVCNLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQLTKKVNNIETGVALGATFHL 420

Qy 421 LQSLGISH 428
Db 421 LQSLGISH 428

RESULT 12
US-09-370-265-5
; Sequence 5, Application US/09370265
; Patent No. 644771
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; APPLICANT: Mulero, Julio
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE
; FILE REFERENCE: 28111/35908
; CURRENT APPLICATION NUMBER: US/09/370,265
; CURRENT FILING DATE: 1999-08-09
; EARLIER APPLICATION NUMBER: PCT/US99/16180
; EARLIER FILING DATE: 1999-07-16
; EARLIER APPLICATION NUMBER: 09/350,836
; EARLIER FILING DATE: 1999-07-09
; EARLIER APPLICATION NUMBER: 09/273,447
; EARLIER FILING DATE: 1999-03-19
; EARLIER APPLICATION NUMBER: 09/350,836
; EARLIER FILING DATE: 1999-07-09
; EARLIER APPLICATION NUMBER: 09/244,444
; EARLIER FILING DATE: 1999-02-04
; EARLIER APPLICATION NUMBER: 09/122,449
; EARLIER FILING DATE: 1998-07-24
; EARLIER APPLICATION NUMBER: 09/118,205
; EARLIER FILING DATE: 1998-07-16
; NUMBER OF SEQ ID NOS: 37
```



; PRIOR FILING DATE: 1998-07-16  
; NUMBER OF SEQ ID NOS: 56  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 5  
; LENGTH: 428  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-557-800C-5

Query Match 99.2%; Score 2235; DB 4; Length 428;  
Best Local Similarity 99.3%; Pred. No. 3.7e-245;  
Matches 425; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
Qy 1 MATSWGTVFVPMVVCVCSAVSHRNQQTWPEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60  
Db 1 MATSWGTVFVPMVVCVCSAVSHRNQQTWPEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60  
Qy 61 RIHYVTVQKMPGOLPILGEVFDVSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120  
Db 61 RIHYVTVQKMPGOLPILGEVFDVSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120  
Qy 121 KKTVPVLKATAGLRLPEHKAKALLFEVKEIFRKSPLVPKGSVSIIMTGDGEGIFAWTV 180  
Db 121 KKTVPVLKATAGLRLPEHKAKALLFEVKEIFRKSPLVPKGSVSIIMTGDGEGIFAWTV 180  
Qy 181 NFLTQGLHGRQETVGTDLGGASTQITFLPQPEKTLQTPRGYLTSEFMFNSTYKLYTH 240  
Db 181 NFLTQGLHGRQETVGTDLGGASTQITFLPQPEKTLQTPRGYLTSEFMFNSTYKLYTH 240  
Qy 241 SYLGFGKKAARLATLGALETGHTGHTFRSACLPRWLEAEWIFGVKQYQYGGNQEGBVGF 300  
Db 241 SYLGFGKKAARLATLGALETGHTGHTFRSACLPRWLEAEWIFGVKQYQYGGNQEGBVGF 300  
Qy 301 EPCYAEVLRVVRGKLHQPVEEVQVRSFYAFSYVDRAVDTMDIDYKGGILKVEDFERKAR 360  
Db 301 EPCYAEVLRVVRGKLHQPVEEVQVRSFYAFSYVDRAVDTMDIDYKGGILKVEDFERKAR 360  
Qy 361 EVCDNLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQLTQTKVNNIETGVALGATFHL 420  
Db 361 EVCDNLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQLTQTKVNNIETGVALGATFHL 420  
Qy 421 LQSLGISH 428  
Db 421 LQSLGISH 428

## RESULT 15

US-09-608-285A-25  
; Sequence 25, Application US/09608285A  
; Patent No. 6335013  
; GENERAL INFORMATION:  
; APPLICANT: Ford, John  
; APPLICANT: Mulero, Julio  
; APPLICANT: Yeung, George  
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO CD39-LIKE  
; FILE REFERENCE: 28110/36570  
; CURRENT APPLICATION NUMBER: US/09/608,285A  
; CURRENT FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: 09/583,231  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 09/557,800  
; PRIOR FILING DATE: 2000-04-25  
; PRIOR APPLICATION NUMBER: 09/481,238  
; PRIOR FILING DATE: 2000-01-11  
; PRIOR APPLICATION NUMBER: 09/370,265  
; PRIOR FILING DATE: 1999-08-09  
; PRIOR APPLICATION NUMBER: PCT/US99/16180  
; PRIOR FILING DATE: 1999-07-16  
; PRIOR APPLICATION NUMBER: 09/350,836  
; PRIOR FILING DATE: 1999-07-09  
; PRIOR APPLICATION NUMBER: 09/273,447  
; PRIOR FILING DATE: 1999-03-19

; PRIOR APPLICATION NUMBER: 09/244,444  
; PRIOR FILING DATE: 1999-02-04  
; PRIOR APPLICATION NUMBER: 09/122,449  
; PRIOR FILING DATE: 1998-07-24  
; PRIOR APPLICATION NUMBER: 09/118,205  
; PRIOR FILING DATE: 1998-07-16  
; NUMBER OF SEQ ID NOS: 60  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 25  
; LENGTH: 405  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-608-285A-25

Query Match 92.8%; Score 2089; DB 4; Length 405;  
Best Local Similarity 99.2%; Pred. No. 1.4e-228;  
Matches 397; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
Qy 1 MATSWGTVFVPMVVCVCSAVSHRNQQTWPEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60  
Db 1 MATSWGTVFVPMVVCVCSAVSHRNQQTWPEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60  
Qy 61 RIHYVTVQKMPGOLPILGEVFDVSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120  
Db 61 RIHYVTVQKMPGOLPILGEVFDVSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120  
Qy 121 KKTVPVLKATAGLRLPEHKAKALLFEVKEIFRKSPLVPKGSVSIIMTGDGEGIFAWTV 180  
Db 121 KKTVPVLKATAGLRLPEHKAKALLFEVKEIFRKSPLVPKGSVSIIMTGDGEGIFAWTV 180  
Qy 181 NFLTQGLHGRQETVGTDLGGASTQITFLPQPEKTLQTPRGYLTSEFMFNSTYKLYTH 240  
Db 181 NFLTQGLHGRQETVGTDLGGASTQITFLPQPEKTLQTPRGYLTSEFMFNSTYKLYTH 240  
Qy 241 SYLGFGKKAARLATLGALETGHTGHTFRSACLPRWLEAEWIFGVKQYQYGGNQEGBVGF 300  
Db 241 SYLGFGKKAARLATLGALETGHTGHTFRSACLPRWLEAEWIFGVKQYQYGGNQEGBVGF 300  
Qy 301 EPCYAEVLRVVRGKLHQPVEEVQVRSFYAFSYVDRAVDTMDIDYKGGILKVEDFERKAR 360  
Db 301 EPCYAEVLRVVRGKLHQPVEEVQVRSFYAFSYVDRAVDTMDIDYKGGILKVEDFERKAR 360  
Qy 361 EVCDNLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQ 400  
Db 361 EVCDNLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQ 400

Search completed: January 29, 2004, 10:06:58  
Job time : 16 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 10:02:15 ; Search time 15 Seconds  
(without alignments)  
1207.270 Million cell updates/sec

Title: US-10-091-085-3  
Perfect score: 2250  
Sequence: 1 MATSNGTFFMLVSVCSA.....ETGALGATFHLQLSLGISH 428

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA: \*  
1: /cgn2\_6/prodata/1/iaa/5A COMB.pep.\*  
2: /cgn2\_6/prodata/1/iaa/5B COMB.pep.\*  
3: /cgn2\_6/prodata/1/iaa/6A COMB.pep.\*  
4: /cgn2\_6/prodata/1/iaa/6B COMB.pep.\*  
5: /cgn2\_6/prodata/1/iaa/PCTUS COMB.pep.\*  
6: /cgn2\_6/prodata/1/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2250	100.0	428	4	US-09-608-285A-3
2	2250	100.0	428	4	US-09-608-285A-5
3	2250	100.0	428	4	US-09-240-639-6
4	2250	100.0	428	4	US-09-240-639-9
5	2250	100.0	428	4	US-09-350-8368-3
6	2250	100.0	428	4	US-09-350-8368-5
7	2250	100.0	428	4	US-09-370-265-3
8	2250	100.0	428	4	US-09-370-265-5
9	2250	100.0	428	4	US-09-557-800C-3
10	2250	100.0	428	4	US-09-557-800C-5
11	2235	99.3	428	4	US-09-608-285A-7
12	2235	99.3	428	4	US-09-350-8368-7
13	2235	99.3	428	4	US-09-370-265-7
14	2235	99.3	428	4	US-09-557-800C-7
15	2104	93.5	405	4	US-09-608-285A-25
16	2104	93.5	405	4	US-09-370-265-25
17	2104	93.5	405	4	US-09-557-800C-25
18	1837.5	81.7	465	4	US-09-557-800C-56
19	1832.5	81.4	465	4	US-09-240-639-8
20	999	44.4	456	4	US-09-240-639-2
21	999	44.4	484	4	US-09-608-285A-27
22	999	44.4	484	4	US-09-370-265-27
23	999	44.4	484	4	US-09-557-800C-27
24	819.5	36.4	471	4	US-09-608-285A-60
25	507.5	22.6	459	4	US-09-129-112-9
26	503.5	22.4	467	4	US-09-129-112-19
27	498	22.1	462	4	US-09-129-112-2

28	491.5	21.8	462	4	US-09-129-112-15	Sequence 15, Appl
29	489	21.7	455	4	US-09-240-639-10	Sequence 10, Appl
30	467.5	20.8	473	4	US-09-240-639-12	Sequence 12, Appl
31	446.5	19.8	454	4	US-09-240-639-11	Sequence 11, Appl
32	351	15.6	502	4	US-09-557-800C-55	Sequence 55, Appl
33	351	15.6	510	3	US-08-930-921-1	Sequence 1, Appl
34	347.5	15.4	529	4	US-09-240-639-4	Sequence 4, Appl
35	234	10.4	148	4	US-09-240-639-17	Sequence 17, Appl
36	186	8.3	153	4	US-09-240-639-13	Sequence 13, Appl
37	180.5	8.0	150	4	US-09-240-639-16	Sequence 16, Appl
38	179	8.0	153	4	US-09-240-639-15	Sequence 15, Appl
39	174	7.7	154	4	US-09-240-639-14	Sequence 14, Appl
40	106.5	4.7	553	1	US-08-565-386-12	Sequence 12, Appl
41	105.5	4.7	339	4	US-09-252-991A-20854	Sequence 20854, A
42	98.5	4.4	535	3	US-09-137-077-2	Sequence 2, Appl
43	96	4.3	506	4	US-09-134-001C-4383	Sequence 4383, Ap
44	93	4.1	969	2	US-08-284-941-2	Sequence 2, Appl
45	93	4.1	969	2	US-08-447-642-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1  
US-09-608-285A-3  
; Sequence 3, Application US/09608285A  
; Patent No. 6335013  
; GENERAL INFORMATION:  
; APPLICANT: Ford, John  
; APPLICANT: Mulero, Julio  
; APPLICANT: Yeung, George  
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO CD39-LIKE  
; FILE REFERENCE: 28110/36570  
; CURRENT APPLICATION NUMBER: US/09/608,285A  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: 09/583,231  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 09/557,800  
; PRIOR FILING DATE: 2000-04-25  
; PRIOR APPLICATION NUMBER: 09/481,238  
; PRIOR FILING DATE: 2000-01-11  
; PRIOR APPLICATION NUMBER: 09/370,265  
; PRIOR FILING DATE: 1999-08-09  
; PRIOR APPLICATION NUMBER: PCT/US99/16180  
; PRIOR FILING DATE: 1999-07-16  
; PRIOR APPLICATION NUMBER: 09/350,836  
; PRIOR FILING DATE: 1999-07-09  
; PRIOR APPLICATION NUMBER: 09/273,447  
; PRIOR FILING DATE: 1999-03-19  
; PRIOR APPLICATION NUMBER: 09/244,444  
; PRIOR FILING DATE: 1999-02-04  
; PRIOR APPLICATION NUMBER: 09/122,449  
; PRIOR FILING DATE: 1998-07-24  
; PRIOR APPLICATION NUMBER: 09/118,205  
; PRIOR FILING DATE: 1998-07-16  
; NUMBER OF SEQ ID NOS: 60  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 428  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-608-285A-3

Query Match 100.0%; Score 2250; DB 4; Length 428;  
Best Local Similarity 100.0%; Pred. No. 3.9e-248;  
Matches 428; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MATSNGTFFMLVSVCSAVSHRNQOTWFEGLFLSSMCPINVSASTLYGIMFDAGSTGT 60  
Db 1 MATSNGTFFMLVSVCSAVSHRNQOTWFEGLFLSSMCPINVSASTLYGIMFDAGSTGT 60  
QY 61 RIHVYTFVQKMPQLFILEGEVDSVKPGLSAFVDQFKOGAETVQGLLEVAKDSIPRSHW 120

Db 61 RIHVYTFVQKMPGOLPILGEVFDVSKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120  
Qy 121 KKTVPVLKATAGLRLLEPEHAKALLFEVKEIFRKSPLVPKGSVIMDGSDEGILAWTV 180  
Db 121 KKTVPVLKATAGLRLLEPEHAKALLFEVKEIFRKSPLVPKGSVIMDGSDEGILAWTV 180  
Qy 181 NFLTQGLHGRQETVGTDLGGASTQITFLPQPEKTELEQTPRGYLTSEFNFNSTYKLYTH 240  
Db 181 NFLTQGLHGRQETVGTDLGGASTQITFLPQPEKTELEQTPRGYLTSEFNFNSTYKLYTH 240  
Qy 241 SYLGFGGLKAARLATLGALETGTDGHTFRSACLPRWLEAEWIFGGVXYQYGGNOEGEVGF 300  
Db 241 SYLGFGGLKAARLATLGALETGTDGHTFRSACLPRWLEAEWIFGGVXYQYGGNOEGEVGF 300  
Qy 301 EPCVAEVLVRVVRGKLHQPVEEVQVRSFYAFSYYVDRAVDTMDIDYEKGGLKVEDFERKAR 360  
Db 301 EPCVAEVLVRVVRGKLHQPVEEVQVRSFYAFSYYVDRAVDTMDIDYEKGGLKVEDFERKAR 360  
Qy 361 EVCNLENFTSGSPFLCMLSYITALLKDGFGFADSTVLQLTKKVNNIETGVALGATFHL 420  
Db 361 EVCNLENFTSGSPFLCMLSYITALLKDGFGFADSTVLQLTKKVNNIETGVALGATFHL 420  
Qy 421 LOSLGISH 428  
Db 421 LOSLGISH 428

RESULT 2  
US-09-608-285A-5  
; Sequence 5, Application US/09608285A  
; Patent No. 6335013  
; GENERAL INFORMATION:  
; APPLICANT: Ford, John  
; APPLICANT: Mulero, Julio  
; APPLICANT: Yeung, George  
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO CD39-LIKE  
; FILE REFERENCE: 28110/36570  
; CURRENT APPLICATION NUMBER: US/09/608,285A  
; PRIOR FILING DATE: 2000-06-10  
; PRIOR APPLICATION NUMBER: 09/583,231  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 09/557,800  
; PRIOR FILING DATE: 2000-04-25  
; PRIOR APPLICATION NUMBER: 09/481,238  
; PRIOR FILING DATE: 2000-01-11  
; PRIOR APPLICATION NUMBER: 09/370,265  
; PRIOR FILING DATE: 1999-08-09  
; PRIOR APPLICATION NUMBER: PCT/US99/16180  
; PRIOR FILING DATE: 1999-07-16  
; PRIOR APPLICATION NUMBER: 09/350,836  
; PRIOR FILING DATE: 1999-07-09  
; PRIOR APPLICATION NUMBER: 09/273,447  
; PRIOR FILING DATE: 1999-03-19  
; PRIOR APPLICATION NUMBER: 09/244,444  
; PRIOR FILING DATE: 1999-02-04  
; PRIOR APPLICATION NUMBER: 09/122,449  
; PRIOR FILING DATE: 1998-07-24  
; PRIOR APPLICATION NUMBER: 09/118,205  
; PRIOR FILING DATE: 1998-07-16  
; NUMBER OF SEQ ID NOS: 60  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 5  
; LENGTH: 428  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-608-285A-5

Query Match 100.0%; Score 2250; DB 4; Length 428;  
Best Local Similarity 100.0%; Pred. No. 3,9e-248;  
Matches 428; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MATSWGTVFPMVLVVCVCSAVSHRNOQTWPEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60  
Db 1 MATSWGTVFPMVLVVCVCSAVSHRNOQTWPEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60  
Qy 61 RIHVYTFVQKMPGOLPILGEVFDVSKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120  
Db 61 RIHVYTFVQKMPGOLPILGEVFDVSKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120  
Qy 121 KKTVPVLKATAGLRLLEPEHAKALLFEVKEIFRKSPLVPKGSVIMDGSDEGILAWTV 180  
Db 121 KKTVPVLKATAGLRLLEPEHAKALLFEVKEIFRKSPLVPKGSVIMDGSDEGILAWTV 180  
Qy 181 NFLTQGLHGRQETVGTDLGGASTQITFLPQPEKTELEQTPRGYLTSEFNFNSTYKLYTH 240  
Db 181 NFLTQGLHGRQETVGTDLGGASTQITFLPQPEKTELEQTPRGYLTSEFNFNSTYKLYTH 240  
Qy 241 SYLGFGGLKAARLATLGALETGTDGHTFRSACLPRWLEAEWIFGGVXYQYGGNOEGEVGF 300  
Db 241 SYLGFGGLKAARLATLGALETGTDGHTFRSACLPRWLEAEWIFGGVXYQYGGNOEGEVGF 300  
Qy 301 EPCVAEVLVRVVRGKLHQPVEEVQVRSFYAFSYYVDRAVDTMDIDYEKGGLKVEDFERKAR 360  
Db 301 EPCVAEVLVRVVRGKLHQPVEEVQVRSFYAFSYYVDRAVDTMDIDYEKGGLKVEDFERKAR 360  
Qy 361 EVCNLENFTSGSPFLCMLSYITALLKDGFGFADSTVLQLTKKVNNIETGVALGATFHL 420  
Db 361 EVCNLENFTSGSPFLCMLSYITALLKDGFGFADSTVLQLTKKVNNIETGVALGATFHL 420  
Qy 421 LOSLGISH 428  
Db 421 LOSLGISH 428

RESULT 3  
US-09-240-639-6  
; Sequence 6, Application US/09240639  
; Patent No. 6350447  
; GENERAL INFORMATION:  
; APPLICANT: Chadwick, Brian Paul  
; APPLICANT: Frischauf, Anna-Maria  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO CD39-LIKE  
; FILE REFERENCE: 9598-066  
; CURRENT APPLICATION NUMBER: US/09/240,639  
; CURRENT FILING DATE: 1998-01-29  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 6  
; LENGTH: 428  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-240-639-6

Query Match 100.0%; Score 2250; DB 4; Length 428;  
Best Local Similarity 100.0%; Pred. No. 3,9e-248;  
Matches 428; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MATSWGTVFPMVLVVCVCSAVSHRNOQTWPEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60  
Db 1 MATSWGTVFPMVLVVCVCSAVSHRNOQTWPEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60  
Qy 61 RIHVYTFVQKMPGOLPILGEVFDVSKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120  
Db 61 RIHVYTFVQKMPGOLPILGEVFDVSKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120  
Qy 121 KKTVPVLKATAGLRLLEPEHAKALLFEVKEIFRKSPLVPKGSVIMDGSDEGILAWTV 180  
Db 121 KKTVPVLKATAGLRLLEPEHAKALLFEVKEIFRKSPLVPKGSVIMDGSDEGILAWTV 180  
Qy 181 NFLTQGLHGRQETVGTDLGGASTQITFLPQPEKTELEQTPRGYLTSEFNFNSTYKLYTH 240  
Db 181 NFLTQGLHGRQETVGTDLGGASTQITFLPQPEKTELEQTPRGYLTSEFNFNSTYKLYTH 240

Query Match 100.0%; Score 2250; DB 4; Length 428;  
Best Local Similarity 100.0%; Pred. No. 3,9e-248;  
Matches 428; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CD39L2 polyphid

QY 241 SYLGFLKAARLATLGALETGTGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEVEVF 300  
 DB 241 SYLGFLKAARLATLGALETGTGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEVEVF 300  
 QY 301 EPCYAEVLVRVGRKLHQPVEVQSGSFYAFSYDDRAVDMDIDYKGGILKVEDFERKAR 360  
 DB 301 EPCYAEVLVRVGRKLHQPVEVQSGSFYAFSYDDRAVDMDIDYKGGILKVEDFERKAR 360  
 QY 361 EVCDNLNFTSGSPFLCMLDSYITALLKDGFGFADSTVLQTLTKKVNNIETGALGATFHL 420  
 DB 361 EVCDNLNFTSGSPFLCMLDSYITALLKDGFGFADSTVLQTLTKKVNNIETGALGATFHL 420  
 QY 421 LQSLGISH 428  
 DB 421 LQSLGISH 428

RESULT 4  
 US-09-240-639-9  
 ; Sequence 9, Application US/09240639  
 ; Patent No. 6350447  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Chadwick, Brian Paul  
 ; APPLICANT: Frischauf, Anna-Maria  
 ; TITLE OF INVENTION: METHODS AND COMPOSITIONS RELATING TO CD39-LIKE  
 ; TITLE OF INVENTION: POLYPEPTIDES AND NUCLEIC ACIDS  
 ; FILE REFERENCE: 9598-066  
 ; CURRENT APPLICATION NUMBER: US/09/240,639  
 ; CURRENT FILING DATE: 1998-01-29  
 ; NUMBER OF SEQ ID NOS: 29  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 9  
 ; LENGTH: 428  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-240-639-9

Query Match 100.0%; Score 2250; DB 4; Length 428;  
 Best Local Similarity 100.0%; Pred. No. 3.9e-248;  
 Matches 428; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATSWGTVFVFMVSVCSAVSHRNQOTWTFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60  
 DB 1 MATSWGTVFVFMVSVCSAVSHRNQOTWTFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60  
 QY 61 RIHVYTFVQKMPQQLPILGEVDFSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120  
 DB 61 RIHVYTFVQKMPQQLPILGEVDFSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120  
 QY 121 KKTVPVVKATAGLRLPEHKAKALLPEVKIIFRKSFLPKGVSIMDGSDEGILAWTV 180  
 DB 121 KKTVPVVKATAGLRLPEHKAKALLPEVKIIFRKSFLPKGVSIMDGSDEGILAWTV 180  
 QY 181 NFLTGQHGHRQETVGTLDLGGASTQITFLPQPEKTLQTPRGYLTSEFNFNSTYKLYTH 240  
 DB 181 NFLTGQHGHRQETVGTLDLGGASTQITFLPQPEKTLQTPRGYLTSEFNFNSTYKLYTH 240  
 QY 241 SYLGFLKAARLATLGALETGTGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEVEVF 300  
 DB 241 SYLGFLKAARLATLGALETGTGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEVEVF 300  
 QY 301 EPCYAEVLVRVGRKLHQPVEVQSGSFYAFSYDDRAVDMDIDYKGGILKVEDFERKAR 360  
 DB 301 EPCYAEVLVRVGRKLHQPVEVQSGSFYAFSYDDRAVDMDIDYKGGILKVEDFERKAR 360  
 QY 361 EVCDNLNFTSGSPFLCMLDSYITALLKDGFGFADSTVLQTLTKKVNNIETGALGATFHL 420  
 DB 361 EVCDNLNFTSGSPFLCMLDSYITALLKDGFGFADSTVLQTLTKKVNNIETGALGATFHL 420  
 QY 421 LQSLGISH 428  
 DB 421 LQSLGISH 428

RESULT 5  
 US-09-350-836B-3  
 ; Sequence 3, Application US/09350836B  
 ; Patent No. 6387645  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ford, John  
 ; APPLICANT: Mulero, Julio  
 ; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE  
 ; TITLE OF INVENTION: POLYPEPTIDES  
 ; FILE REFERENCE: 28110/35761  
 ; CURRENT APPLICATION NUMBER: US/09/350,836B  
 ; CURRENT FILING DATE: 1999-07-09  
 ; PRIOR APPLICATION NUMBER: 09/273,447  
 ; PRIOR FILING DATE: 1999-03-19  
 ; PRIOR APPLICATION NUMBER: 09/118,205  
 ; PRIOR FILING DATE: 1998-07-16  
 ; PRIOR APPLICATION NUMBER: 09/122,449  
 ; PRIOR FILING DATE: 1998-07-24  
 ; PRIOR APPLICATION NUMBER: 09/244,444  
 ; PRIOR FILING DATE: 1999-02-04  
 ; NUMBER OF SEQ ID NOS: 23  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 3  
 ; LENGTH: 428  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-350-836B-3

Query Match 100.0%; Score 2250; DB 4; Length 428;  
 Best Local Similarity 100.0%; Pred. No. 3.9e-248;  
 Matches 428; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATSWGTVFVFMVSVCSAVSHRNQOTWTFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60  
 DB 1 MATSWGTVFVFMVSVCSAVSHRNQOTWTFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60  
 QY 61 RIHVYTFVQKMPQQLPILGEVDFSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120  
 DB 61 RIHVYTFVQKMPQQLPILGEVDFSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120  
 QY 121 KKTVPVVKATAGLRLPEHKAKALLPEVKIIFRKSFLPKGVSIMDGSDEGILAWTV 180  
 DB 121 KKTVPVVKATAGLRLPEHKAKALLPEVKIIFRKSFLPKGVSIMDGSDEGILAWTV 180  
 QY 181 NFLTGQHGHRQETVGTLDLGGASTQITFLPQPEKTLQTPRGYLTSEFNFNSTYKLYTH 240  
 DB 181 NFLTGQHGHRQETVGTLDLGGASTQITFLPQPEKTLQTPRGYLTSEFNFNSTYKLYTH 240  
 QY 241 SYLGFLKAARLATLGALETGTGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEVEVF 300  
 DB 241 SYLGFLKAARLATLGALETGTGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEVEVF 300  
 QY 301 EPCYAEVLVRVGRKLHQPVEVQSGSFYAFSYDDRAVDMDIDYKGGILKVEDFERKAR 360  
 DB 301 EPCYAEVLVRVGRKLHQPVEVQSGSFYAFSYDDRAVDMDIDYKGGILKVEDFERKAR 360  
 QY 361 EVCDNLNFTSGSPFLCMLDSYITALLKDGFGFADSTVLQTLTKKVNNIETGALGATFHL 420  
 DB 361 EVCDNLNFTSGSPFLCMLDSYITALLKDGFGFADSTVLQTLTKKVNNIETGALGATFHL 420  
 QY 421 LQSLGISH 428  
 DB 421 LQSLGISH 428

RESULT 6  
 US-09-350-836B-5  
 ; Sequence 5, Application US/09350836B  
 ; Patent No. 6387645  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ford, John  
 ; APPLICANT: Mulero, Julio

; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE
; FILE REFERENCE: 28110/35761
; CURRENT APPLICATION NUMBER: US/09/350,836B
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/273,447
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: 09/118,205
; PRIOR FILING DATE: 1998-07-16
; PRIOR APPLICATION NUMBER: 09/122,449
; PRIOR FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: 09/244,444
; PRIOR FILING DATE: 1999-02-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-350-836B-5

Query Match 100.0%; Score 2250; DB 4; Length 428;
Best Local Similarity 100.0%; Pred. No. 3.9e-248;
Matches 428; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MATSGTFFVFMVWSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
Db 1 MATSGTFFVFMVWSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

Qy 61 RIHVYTFVQKMPGQLPILLEGVDSVKPGLSAFVDQPKQGAETVQGLLEAKDSIPRSHW 120
Db 61 RIHVYTFVQKMPGQLPILLEGVDSVKPGLSAFVDQPKQGAETVQGLLEAKDSIPRSHW 120

Qy 121 KKTVPVLKATAGLRLPEHKAKALLFEVKEIFRKSPLVPKGSVIMDGSDEGILAWTV 180
Db 121 KKTVPVLKATAGLRLPEHKAKALLFEVKEIFRKSPLVPKGSVIMDGSDEGILAWTV 180

Qy 181 NFLTQGLHGRQETVGTLDLGGASTQITFLPQPEKTLQTPRGYLTSPFEMENSTYKLYTH 240
Db 181 NFLTQGLHGRQETVGTLDLGGASTQITFLPQPEKTLQTPRGYLTSPFEMENSTYKLYTH 240

Qy 241 SYLGFGLKAARLATLGALETGTGHTFRSACLPRWLEAEWIFGVKYQYGGNOGEVGF 300
Db 241 SYLGFGLKAARLATLGALETGTGHTFRSACLPRWLEAEWIFGVKYQYGGNOGEVGF 300

Qy 301 EPCVAEVLVRVGRKQHPEEVQSGSFYAFSYYYDRAVDTMDIDYKGGILKVEDFERKAR 360
Db 301 EPCVAEVLVRVGRKQHPEEVQSGSFYAFSYYYDRAVDTMDIDYKGGILKVEDFERKAR 360

Qy 361 EVCNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVQLTKKVNNIETGALGATFHL 420
Db 361 EVCNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVQLTKKVNNIETGALGATFHL 420

Qy 421 LQSLGISH 428
Db 421 LQSLGISH 428

RESULT 7
US-09-370-265-3
; Sequence 3, Application US/09370265
; Patent No. 6447771
; GENERAL INFORMATION:
; APPLICANT: Mulero, Julio
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE
; FILE REFERENCE: 28111/35908
; CURRENT APPLICATION NUMBER: US/09/370,265
; CURRENT FILING DATE: 1999-08-09
; EARLIER APPLICATION NUMBER: PCT/US99/16180
; EARLIER FILING DATE: 1999-07-16
; EARLIER APPLICATION NUMBER: 09/350,836
; EARLIER FILING DATE: 1999-03-19
; EARLIER APPLICATION NUMBER: 09/273,447
; EARLIER FILING DATE: 1999-03-19
; EARLIER APPLICATION NUMBER: 09/244,444

; EARLIER FILING DATE: 1999-07-09
; EARLIER APPLICATION NUMBER: 09/273,447
; EARLIER FILING DATE: 1999-03-19
; EARLIER APPLICATION NUMBER: 09/244,444
; EARLIER FILING DATE: 1999-02-04
; EARLIER APPLICATION NUMBER: 09/122,449
; EARLIER FILING DATE: 1998-07-24
; EARLIER APPLICATION NUMBER: 09/118,205
; EARLIER FILING DATE: 1998-07-16
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-370-265-3

Query Match 100.0%; Score 2250; DB 4; Length 428;
Best Local Similarity 100.0%; Pred. No. 3.9e-248;
Matches 428; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MATSGTFFVFMVWSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
Db 1 MATSGTFFVFMVWSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

Qy 61 RIHVYTFVQKMPGQLPILLEGVDSVKPGLSAFVDQPKQGAETVQGLLEAKDSIPRSHW 120
Db 61 RIHVYTFVQKMPGQLPILLEGVDSVKPGLSAFVDQPKQGAETVQGLLEAKDSIPRSHW 120

Qy 121 KKTVPVLKATAGLRLPEHKAKALLFEVKEIFRKSPLVPKGSVIMDGSDEGILAWTV 180
Db 121 KKTVPVLKATAGLRLPEHKAKALLFEVKEIFRKSPLVPKGSVIMDGSDEGILAWTV 180

Qy 181 NFLTQGLHGRQETVGTLDLGGASTQITFLPQPEKTLQTPRGYLTSPFEMENSTYKLYTH 240
Db 181 NFLTQGLHGRQETVGTLDLGGASTQITFLPQPEKTLQTPRGYLTSPFEMENSTYKLYTH 240

Qy 241 SYLGFGLKAARLATLGALETGTGHTFRSACLPRWLEAEWIFGVKYQYGGNOGEVGF 300
Db 241 SYLGFGLKAARLATLGALETGTGHTFRSACLPRWLEAEWIFGVKYQYGGNOGEVGF 300

Qy 301 EPCVAEVLVRVGRKQHPEEVQSGSFYAFSYYYDRAVDTMDIDYKGGILKVEDFERKAR 360
Db 301 EPCVAEVLVRVGRKQHPEEVQSGSFYAFSYYYDRAVDTMDIDYKGGILKVEDFERKAR 360

Qy 361 EVCNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVQLTKKVNNIETGALGATFHL 420
Db 361 EVCNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVQLTKKVNNIETGALGATFHL 420

Qy 421 LQSLGISH 428
Db 421 LQSLGISH 428

RESULT 8
US-09-370-265-5
; Sequence 5, Application US/09370265
; Patent No. 6447771
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; APPLICANT: Mulero, Julio
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE
; FILE REFERENCE: 28111/35908
; CURRENT APPLICATION NUMBER: US/09/370,265
; CURRENT FILING DATE: 1999-08-09
; EARLIER APPLICATION NUMBER: PCT/US99/16180
; EARLIER FILING DATE: 1999-07-16
; EARLIER APPLICATION NUMBER: 09/350,836
; EARLIER FILING DATE: 1999-07-09
; EARLIER APPLICATION NUMBER: 09/273,447
; EARLIER FILING DATE: 1999-03-19
; EARLIER APPLICATION NUMBER: 09/244,444

EARLIER FILING DATE: 1999-02-04  
; EARLIER APPLICATION NUMBER: 09/122,449  
; EARLIER FILING DATE: 1998-07-24  
; EARLIER APPLICATION NUMBER: 09/244,444  
; EARLIER FILING DATE: 1999-02-04  
; EARLIER APPLICATION NUMBER: 09/118,205  
; EARLIER FILING DATE: 1998-07-16  
; NUMBER OF SEQ ID NOS: 37  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 5  
; LENGTH: 428  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-370-265-5

Query Match 100.0%; Score 2250; DB 4; Length 428;  
Best Local Similarity 100.0%; Pred. No. 3.9e-248;  
Matches 428; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATSWGTVPFVLMVSVCSAVSHRNQOTWFEIGIFLSSMCPINVSASTLYGIMFDAGSTGT 60  
DB 1 MATSWGTVPFVLMVSVCSAVSHRNQOTWFEIGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

QY 61 RIHVYTFVQKMPGQLPILGEVDFSVKPGLSAFVDPQKGAETVQGLLEVAKDSIPRSHW 120  
DB 61 RIHVYTFVQKMPGQLPILGEVDFSVKPGLSAFVDPQKGAETVQGLLEVAKDSIPRSHW 120

QY 121 KKTVPVVKATAGLRLLPEHKAKALLFEVKEIFRKSPPFLVPKGSVIMDGSDEGILAWTV 180  
DB 121 KKTVPVVKATAGLRLLPEHKAKALLFEVKEIFRKSPPFLVPKGSVIMDGSDEGILAWTV 180

QY 181 NFLTGQHGHRQETVGTDLGGASTQITFLPQPEKLTLEOTPRGYLTSFEMFNSTYKLYTH 240  
DB 181 NFLTGQHGHRQETVGTDLGGASTQITFLPQPEKLTLEOTPRGYLTSFEMFNSTYKLYTH 240

QY 241 SYLGFGKKAARLATLGALETGDTGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEVEVGF 300  
DB 241 SYLGFGKKAARLATLGALETGDTGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEVEVGF 300

QY 301 EPCYAEVLRVVRGKLHQPVEVQSGSFYAFSYYVDRAVDTMDIDYEKGGILKVEDFERKAR 360  
DB 301 EPCYAEVLRVVRGKLHQPVEVQSGSFYAFSYYVDRAVDTMDIDYEKGGILKVEDFERKAR 360

QY 361 EVCNLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQLTKKVNNIETGVALGATFHL 420  
DB 361 EVCNLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQLTKKVNNIETGVALGATFHL 420

QY 421 LQSLGISH 428  
DB 421 LQSLGISH 428

RESULT 9  
US-09-557-800C-3  
; Sequence 3, Application US/09557800C  
; Patent No. 6476211  
; GENERAL INFORMATION:  
; APPLICANT: Ford, John  
; APPLICANT: Mulero, Julio  
; APPLICANT: Young, George  
; TITLE OF INVENTION: Methods and Materials Relating to CD39-Like  
; TITLE OF INVENTION: Polypeptides  
; FILE REFERENCE: 28110/36457  
; CURRENT APPLICATION NUMBER: US/09/557,800C  
; CURRENT FILING DATE: 2000-04-25  
; PRIOR APPLICATION NUMBER: 09/481,238  
; PRIOR FILING DATE: 2000-01-11  
; PRIOR APPLICATION NUMBER: 09/370,265  
; PRIOR FILING DATE: 1999-08-09  
; PRIOR APPLICATION NUMBER: PCT/US99/16180  
; PRIOR FILING DATE: 1999-07-16  
; PRIOR APPLICATION NUMBER: 09/350836  
; PRIOR FILING DATE: 1999-07-09  
; PRIOR APPLICATION NUMBER: 09/273447  
; PRIOR FILING DATE: 1999-03-19

PRIOR APPLICATION NUMBER: 09/122449  
; PRIOR FILING DATE: 1998-07-24  
; PRIOR APPLICATION NUMBER: 09/244444  
; PRIOR FILING DATE: 1999-02-04  
; PRIOR APPLICATION NUMBER: 09/118,205  
; PRIOR FILING DATE: 1998-07-16  
; NUMBER OF SEQ ID NOS: 56  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 428  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-557-800C-3

Query Match 100.0%; Score 2250; DB 4; Length 428;  
Best Local Similarity 100.0%; Pred. No. 3.9e-248;  
Matches 428; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATSWGTVPFVLMVSVCSAVSHRNQOTWFEIGIFLSSMCPINVSASTLYGIMFDAGSTGT 60  
DB 1 MATSWGTVPFVLMVSVCSAVSHRNQOTWFEIGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

QY 61 RIHVYTFVQKMPGQLPILGEVDFSVKPGLSAFVDPQKGAETVQGLLEVAKDSIPRSHW 120  
DB 61 RIHVYTFVQKMPGQLPILGEVDFSVKPGLSAFVDPQKGAETVQGLLEVAKDSIPRSHW 120

QY 121 KKTVPVVKATAGLRLLPEHKAKALLFEVKEIFRKSPPFLVPKGSVIMDGSDEGILAWTV 180  
DB 121 KKTVPVVKATAGLRLLPEHKAKALLFEVKEIFRKSPPFLVPKGSVIMDGSDEGILAWTV 180

QY 181 NFLTGQHGHRQETVGTDLGGASTQITFLPQPEKLTLEOTPRGYLTSFEMFNSTYKLYTH 240  
DB 181 NFLTGQHGHRQETVGTDLGGASTQITFLPQPEKLTLEOTPRGYLTSFEMFNSTYKLYTH 240

QY 241 SYLGFGKKAARLATLGALETGDTGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEVEVGF 300  
DB 241 SYLGFGKKAARLATLGALETGDTGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEVEVGF 300

QY 301 EPCYAEVLRVVRGKLHQPVEVQSGSFYAFSYYVDRAVDTMDIDYEKGGILKVEDFERKAR 360  
DB 301 EPCYAEVLRVVRGKLHQPVEVQSGSFYAFSYYVDRAVDTMDIDYEKGGILKVEDFERKAR 360

QY 361 EVCNLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQLTKKVNNIETGVALGATFHL 420  
DB 361 EVCNLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQLTKKVNNIETGVALGATFHL 420

QY 421 LQSLGISH 428  
DB 421 LQSLGISH 428

RESULT 10  
US-09-557-800C-5  
; Sequence 5, Application US/09557800C  
; Patent No. 6476211  
; GENERAL INFORMATION:  
; APPLICANT: Ford, John  
; APPLICANT: Mulero, Julio  
; APPLICANT: Young, George  
; TITLE OF INVENTION: Methods and Materials Relating to CD39-Like  
; TITLE OF INVENTION: Polypeptides  
; FILE REFERENCE: 28110/36457  
; CURRENT APPLICATION NUMBER: US/09/557,800C  
; CURRENT FILING DATE: 2000-04-25  
; PRIOR APPLICATION NUMBER: 09/481,238  
; PRIOR FILING DATE: 2000-01-11  
; PRIOR APPLICATION NUMBER: 09/370,265  
; PRIOR FILING DATE: 1999-08-09  
; PRIOR APPLICATION NUMBER: PCT/US99/16180  
; PRIOR FILING DATE: 1999-07-16  
; PRIOR APPLICATION NUMBER: 09/350836  
; PRIOR FILING DATE: 1999-07-09  
; PRIOR APPLICATION NUMBER: 09/273447

CD 39L 2 Polymultimer

; PRIOR FILING DATE: 1999-03-19  
 ; PRIOR APPLICATION NUMBER: 09/122449  
 ; PRIOR FILING DATE: 1998-07-24  
 ; PRIOR APPLICATION NUMBER: 09/244444  
 ; PRIOR FILING DATE: 1999-02-04  
 ; PRIOR APPLICATION NUMBER: 09/118,205  
 ; PRIOR FILING DATE: 1998-07-16  
 ; NUMBER OF SEQ ID NOS: 56  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 5  
 ; LENGTH: 428  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-557-800C-5

Query Match 100.0%; Score 2250; DB 4; Length 428;  
 Best Local Similarity 100.0%; Pred. No. 3.9e-248; Indels 0; Gaps 0;  
 Matches 428; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 MATSGTGVFFMLVSVCSAVSHRNQQTWTFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60  
 Db 1 MATSGTGVFFMLVSVCSAVSHRNQQTWTFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60  
 Qy 61 RIHYVTFVQKMPGQLPILGEVDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120  
 Db 61 RIHYVTFVQKMPGQLPILGEVDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120  
 Qy 121 KKTVPVLKATAGRLLLPEHKAALLFEVKEIFRKSPLVPKGSVIMDGSDEGILAWTV 180  
 Db 121 KKTVPVLKATAGRLLLPEHKAALLFEVKEIFRKSPLVPKGSVIMDGSDEGILAWTV 180  
 Qy 181 NFLTQLHGRQETVGTLDLGASTQITFLPQPEKTLQTPRGYLTSPFEMFNSTYKLYTH 240  
 Db 181 NFLTQLHGRQETVGTLDLGASTQITFLPQPEKTLQTPRGYLTSPFEMFNSTYKLYTH 240  
 Qy 241 SYLGFGLKAARLATLGALETGTGHTFRSACLPRWLEAEWIFGCVKYQYCGNOGEVGF 300  
 Db 241 SYLGFGLKAARLATLGALETGTGHTFRSACLPRWLEAEWIFGCVKYQYCGNOGEVGF 300  
 Qy 301 EPCVAEVLVRVGRKQHPEEVQVGSFYAFSYVDRAVDTMDIDYKGGILKVEDFERKAR 360  
 Db 301 EPCVAEVLVRVGRKQHPEEVQVGSFYAFSYVDRAVDTMDIDYKGGILKVEDFERKAR 360  
 Qy 361 EVCNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQTLTKVNNIETGALGATFHL 420  
 Db 361 EVCNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQTLTKVNNIETGALGATFHL 420  
 Qy 421 LQSLGISH 428  
 Db 421 LQSLGISH 428

RESULT 11  
 US-09-608-285A-7  
 ; Sequence 7, Application US/09608285A  
 ; Patent No. 6335013  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ford, John  
 ; APPLICANT: Mulero, Julio  
 ; APPLICANT: Yeung, George  
 ; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO CD39-LIKE  
 ; FILE REFERENCE: 28110/36570  
 ; CURRENT APPLICATION NUMBER: US/09/608,285A  
 ; PRIOR APPLICATION NUMBER: 09/583,231  
 ; PRIOR FILING DATE: 2000-05-26  
 ; PRIOR APPLICATION NUMBER: 09/557,800  
 ; PRIOR FILING DATE: 2000-04-25  
 ; PRIOR APPLICATION NUMBER: 09/481,238  
 ; PRIOR FILING DATE: 2000-01-11  
 ; PRIOR APPLICATION NUMBER: 09/370,265  
 ; PRIOR FILING DATE: 1999-08-09

; PRIOR APPLICATION NUMBER: PCT/US99/16180  
 ; PRIOR FILING DATE: 1999-07-16  
 ; PRIOR APPLICATION NUMBER: 09/350,836  
 ; PRIOR FILING DATE: 1999-07-09  
 ; PRIOR APPLICATION NUMBER: 09/273,447  
 ; PRIOR FILING DATE: 1999-03-19  
 ; PRIOR APPLICATION NUMBER: 09/244,444  
 ; PRIOR FILING DATE: 1999-02-04  
 ; PRIOR APPLICATION NUMBER: 09/122,449  
 ; PRIOR FILING DATE: 1998-07-24  
 ; PRIOR APPLICATION NUMBER: 09/118,205  
 ; PRIOR FILING DATE: 1998-07-16  
 ; NUMBER OF SEQ ID NOS: 60  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 7  
 ; LENGTH: 428  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-09-608-285A-7

Query Match 99.3%; Score 2235; DB 4; Length 428;  
 Best Local Similarity 99.3%; Pred. No. 2e-246;  
 Matches 425; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 Qy 1 MATSGTGVFFMLVSVCSAVSHRNQQTWTFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60  
 Db 1 MATSGTGVFFMLVSVCSAVSHRNQQTWTFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60  
 Qy 61 RIHYVTFVQKMPGQLPILGEVDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120  
 Db 61 RIHYVTFVQKMPGQLPILGEVDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120  
 Qy 121 KKTVPVLKATAGRLLLPEHKAALLFEVKEIFRKSPLVPKGSVIMDGSDEGILAWTV 180  
 Db 121 KKTVPVLKATAGRLLLPEHKAALLFEVKEIFRKSPLVPKGSVIMDGSDEGILAWTV 180  
 Qy 181 NFLTQLHGRQETVGTLDLGASTQITFLPQPEKTLQTPRGYLTSPFEMFNSTYKLYTH 240  
 Db 181 NFLTQLHGRQETVGTLDLGASTQITFLPQPEKTLQTPRGYLTSPFEMFNSTYKLYTH 240  
 Qy 241 SYLGFGLKAARLATLGALETGTGHTFRSACLPRWLEAEWIFGCVKYQYCGNOGEVGF 300  
 Db 241 SYLGFGLKAARLATLGALETGTGHTFRSACLPRWLEAEWIFGCVKYQYCGNOGEVGF 300  
 Qy 301 EPCVAEVLVRVGRKQHPEEVQVGSFYAFSYVDRAVDTMDIDYKGGILKVEDFERKAR 360  
 Db 301 EPCVAEVLVRVGRKQHPEEVQVGSFYAFSYVDRAVDTMDIDYKGGILKVEDFERKAR 360  
 Qy 361 EVCNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQTLTKVNNIETGALGATFHL 420  
 Db 361 EVCNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQTLTKVNNIETGALGATFHL 420  
 Qy 421 LQSLGISH 428  
 Db 421 LQSLGISH 428

RESULT 12  
 US-09-350-836B-7  
 ; Sequence 7, Application US/09350836B  
 ; Patent No. 6387645  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ford, John  
 ; APPLICANT: Mulero, Julio  
 ; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE  
 ; FILE REFERENCE: 28110/35761  
 ; CURRENT APPLICATION NUMBER: US/09/350,836B  
 ; PRIOR APPLICATION NUMBER: 09/273,447  
 ; PRIOR FILING DATE: 1999-07-09  
 ; PRIOR APPLICATION NUMBER: 09/118,205  
 ; PRIOR FILING DATE: 1998-07-16

; PRIOR APPLICATION NUMBER: 09/122,449  
; PRIOR FILING DATE: 1998-07-24  
; PRIOR APPLICATION NUMBER: 09/244,444  
; PRIOR FILING DATE: 1999-02-04  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO: 7  
; LENGTH: 428  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-350-836B-7

Query Match 99.3%; Score 2235; DB 4; Length 428;  
Best Local Similarity 99.3%; Pred. No. 2e-246;  
Matches 425; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
  
QY 1 MATSWGTVFVPMVSVCSAVSHRNQOTWFEGLFLSSMCPINVSASTLYGIMFDAGSTGT 60  
DB 1 MATSWGTVFVPMVSVCSAVSHRNQOTWFEGLFLSSMCPINVSASTLYGIMFDAGSTGT 60  
  
QY 61 RIHVYTFVQKMPQQLPILEGVEFDSVKPGLSAFVDPKQGAETVQGLLEVAKDSIPRSHW 120  
DB 61 RIHVYTFVQKMPQQLPILEGVEFDSVKPGLSAFVDPKQGAETVQGLLEVAKDSIPRSHW 120  
  
QY 121 KKTVPVVLKATAGLRLPEHKAKALLFEVKEIFRKSPPFLVPKGSVIMTGDGEGIFAWTV 180  
DB 121 KKTVPVVLKATAGLRLPEHKAKALLFEVKEIFRKSPPFLVPKGSVIMTGDGEGIFAWTV 180  
  
QY 181 NFLTQGLHGHROETVGTDLGGASTQITFLPOFEKTELEQTPRGYLTSEFNFNSTYKLYTH 240  
DB 181 NFLTQGLHGHROETVGTDLGGASTQITFLPOFEKTELEQTPRGYLTSEFNFNSTYKLYTH 240  
  
QY 241 SYLGFGKKAARLATLGALETGTDGHTFRSACLPRMLEAEWIFGGVKYQYGGNQEVEVGF 300  
DB 241 SYLGFGKKAARLATLGALETGTDGHTFRSACLPRMLEAEWIFGGVKYQYGGNQEVEVGF 300  
  
QY 301 EPCYAEVLVRVVRGKLHQPEVQSGSFYAFSYDDRAVDTMDIYKGGILKVEDFERKAR 360  
DB 301 EPCYAEVLVRVVRGKLHQPEVQSGSFYAFSYDDRAVDTMDIYKGGILKVEDFERKAR 360  
  
QY 361 EVCDNLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQTKKVNNIETGVALGATFHL 420  
DB 361 EVCDNLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQTKKVNNIETGVALGATFHL 420  
  
QY 421 LOSLGISH 428  
DB 421 LOSLGISH 428

RESULT 13  
US-09-370-265-7  
; Sequence 7, Application US/09370265  
; Patent No. 644771  
; GENERAL INFORMATION:  
; APPLICANT: Mulero, Julio  
; APPLICANT: Mulero, Julio  
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE  
; FILE REFERENCE: 28111/35908  
; CURRENT APPLICATION NUMBER: US/09/370,265  
; PRIOR FILING DATE: 1999-08-09  
; EARLIER APPLICATION NUMBER: PCT/US99/16180  
; EARLIER FILING DATE: 1999-07-16  
; EARLIER APPLICATION NUMBER: 09/350,836  
; EARLIER FILING DATE: 1999-07-09  
; EARLIER APPLICATION NUMBER: 09/273,447  
; EARLIER FILING DATE: 1999-03-19  
; EARLIER APPLICATION NUMBER: 09/244,444  
; EARLIER FILING DATE: 1999-02-04  
; EARLIER APPLICATION NUMBER: 09/122,449  
; EARLIER FILING DATE: 1998-07-24  
; EARLIER APPLICATION NUMBER: 09/118,205  
; EARLIER FILING DATE: 1998-07-16

; NUMBER OF SEQ ID NOS: 37  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO: 7  
; LENGTH: 428  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-370-265-7

Query Match 99.3%; Score 2235; DB 4; Length 428;  
Best Local Similarity 99.3%; Pred. No. 2e-246;  
Matches 425; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
  
QY 1 MATSWGTVFVPMVSVCSAVSHRNQOTWFEGLFLSSMCPINVSASTLYGIMFDAGSTGT 60  
DB 1 MATSWGTVFVPMVSVCSAVSHRNQOTWFEGLFLSSMCPINVSASTLYGIMFDAGSTGT 60  
  
QY 61 RIHVYTFVQKMPQQLPILEGVEFDSVKPGLSAFVDPKQGAETVQGLLEVAKDSIPRSHW 120  
DB 61 RIHVYTFVQKMPQQLPILEGVEFDSVKPGLSAFVDPKQGAETVQGLLEVAKDSIPRSHW 120  
  
QY 121 KKTVPVVLKATAGLRLPEHKAKALLFEVKEIFRKSPPFLVPKGSVIMTGDGEGIFAWTV 180  
DB 121 KKTVPVVLKATAGLRLPEHKAKALLFEVKEIFRKSPPFLVPKGSVIMTGDGEGIFAWTV 180  
  
QY 181 NFLTQGLHGHROETVGTDLGGASTQITFLPOFEKTELEQTPRGYLTSEFNFNSTYKLYTH 240  
DB 181 NFLTQGLHGHROETVGTDLGGASTQITFLPOFEKTELEQTPRGYLTSEFNFNSTYKLYTH 240  
  
QY 241 SYLGFGKKAARLATLGALETGTDGHTFRSACLPRMLEAEWIFGGVKYQYGGNQEVEVGF 300  
DB 241 SYLGFGKKAARLATLGALETGTDGHTFRSACLPRMLEAEWIFGGVKYQYGGNQEVEVGF 300  
  
QY 301 EPCYAEVLVRVVRGKLHQPEVQSGSFYAFSYDDRAVDTMDIYKGGILKVEDFERKAR 360  
DB 301 EPCYAEVLVRVVRGKLHQPEVQSGSFYAFSYDDRAVDTMDIYKGGILKVEDFERKAR 360  
  
QY 361 EVCDNLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQTKKVNNIETGVALGATFHL 420  
DB 361 EVCDNLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQTKKVNNIETGVALGATFHL 420  
  
QY 421 LOSLGISH 428  
DB 421 LOSLGISH 428

RESULT 14  
US-09-557-800C-7  
; Sequence 7, Application US/09557800C  
; Patent No. 6476211  
; GENERAL INFORMATION:  
; APPLICANT: Mulero, Julio  
; APPLICANT: Mulero, Julio  
; TITLE OF INVENTION: Methods and Materials Relating to CD39-Like  
; FILE REFERENCE: 28110/36457  
; CURRENT APPLICATION NUMBER: US/09/557,800C  
; CURRENT FILING DATE: 2000-04-25  
; PRIOR APPLICATION NUMBER: 09/481,238  
; PRIOR FILING DATE: 2000-01-11  
; PRIOR APPLICATION NUMBER: 09/370,265  
; PRIOR FILING DATE: 1999-08-09  
; PRIOR APPLICATION NUMBER: PCT/US99/16180  
; PRIOR FILING DATE: 1999-07-16  
; PRIOR APPLICATION NUMBER: 09/350836  
; PRIOR FILING DATE: 1999-07-09  
; PRIOR APPLICATION NUMBER: 09/273447  
; PRIOR FILING DATE: 1999-03-19  
; PRIOR APPLICATION NUMBER: 09/122449  
; PRIOR FILING DATE: 1998-07-24  
; PRIOR APPLICATION NUMBER: 09/244444  
; PRIOR FILING DATE: 1999-02-04  
; PRIOR APPLICATION NUMBER: 09/118,205

; PRIOR FILING DATE: 1998-07-16  
 ; NUMBER OF SEQ ID NOS: 56  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 7  
 ; LENGTH: 428  
 ; TYPE: PR1  
 ; ORGANISM: Homo sapiens  
 ; US-09-557-800C-7

Query Match 99.3%; Score 2235; DB 4; Length 428;  
 Best Local Similarity 99.3%; Pred. No. 2e-246;  
 Matches 425; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy	1	MATSWGTFFMLVVCVCSAVSHRNQOTWPEGIFLSSMCPINVSASTLYGIMFDAGSTGT	60
Db	1	MATSWGTFFMLVVCVCSAVSHRNQOTWPEGIFLSSMCPINVSASTLYGIMFDAGSTGT	60
Qy	61	RIHYVTFVQKMPGOLPILGEVFDVSVKPGLSAFVDPKQGAETVQGLLEVAKDSIPRSHW	120
Db	61	RIHYVTFVQKMPGOLPILGEVFDVSVKPGLSAFVDPKQGAETVQGLLEVAKDSIPRSHW	120
Qy	121	KKTPVVLKATAGLRLPEHKAALLFEVKEIFRKSPLVPKGSVIMDGSDEGILAWTV	180
Db	121	KKTPVVLKATAGLRLPEHKAALLFEVKEIFRKSPLVPKGSVIMDGSDEGILAWTV	180
Qy	181	NFLTGQLHGRQETVGTDLGGASTQITFLPQPEKTLQTPRGYLTSEFNFNSTYKLYTH	240
Db	181	NFLTGQLHGRQETVGTDLGGASTQITFLPQPEKTLQTPRGYLTSEFNFNSTYKLYTH	240
Qy	241	SYLGFGLKAARLATLGALETGHTGHTFRSACLPRWLEAEWIFGGVYQYGGNQEVEGF	300
Db	241	SYLGFGLKAARLATLGALETGHTGHTFRSACLPRWLEAEWIFGGVYQYGGNQEVEGF	300
Qy	301	EPCVAEVLVRVGRKHOPEVQVGSFYAFSYVDRAVDTMDIDYKGGILKVEDFERKAR	360
Db	301	EPCVAEVLVRVGRKHOPEVQVGSFYAFSYVDRAVDTMDIDYKGGILKVEDFERKAR	360
Qy	361	EVCDNLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQTKKVNNIETGVALGATFHL	420
Db	361	EVCDNLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQTKKVNNIETGVALGATFHL	420
Qy	421	LQSLGISH 428	
Db	421	LQSLGISH 428	

RESULT 15  
 ; Sequence 25, Application US/09608285A  
 ; Patent No. 6335013  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ford, John  
 ; APPLICANT: Mulero, Julio  
 ; APPLICANT: Yeung, George  
 ; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO CD39-LIKE  
 ; FILE REFERENCE: 28110/36570  
 ; CURRENT APPLICATION NUMBER: US/09/608,285A  
 ; CURRENT FILING DATE: 2000-06-30  
 ; PRIOR APPLICATION NUMBER: 09/583,231  
 ; PRIOR FILING DATE: 2000-05-26  
 ; PRIOR APPLICATION NUMBER: 09/557,800  
 ; PRIOR FILING DATE: 2000-04-25  
 ; PRIOR APPLICATION NUMBER: 09/481,238  
 ; PRIOR FILING DATE: 2000-01-11  
 ; PRIOR APPLICATION NUMBER: 09/370,265  
 ; PRIOR FILING DATE: 1999-08-09  
 ; PRIOR APPLICATION NUMBER: PCT/US99/16180  
 ; PRIOR FILING DATE: 1999-07-16  
 ; PRIOR APPLICATION NUMBER: 09/350,836  
 ; PRIOR FILING DATE: 1999-07-09  
 ; PRIOR APPLICATION NUMBER: 09/273,447  
 ; PRIOR FILING DATE: 1999-03-19

; PRIOR APPLICATION NUMBER: 09/244,444  
 ; PRIOR FILING DATE: 1999-02-04  
 ; PRIOR APPLICATION NUMBER: 09/122,449  
 ; PRIOR FILING DATE: 1998-07-24  
 ; PRIOR APPLICATION NUMBER: 09/118,205  
 ; PRIOR FILING DATE: 1998-07-16  
 ; NUMBER OF SEQ ID NOS: 60  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 25  
 ; LENGTH: 405  
 ; TYPE: PR1  
 ; ORGANISM: Homo sapiens  
 ; US-09-608-285A-25

Query Match 93.5%; Score 2104; DB 4; Length 405;  
 Best Local Similarity 100.0%; Pred. No. 1.7e-231;  
 Matches 400; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MATSWGTFFMLVVCVCSAVSHRNQOTWPEGIFLSSMCPINVSASTLYGIMFDAGSTGT	60
Db	1	MATSWGTFFMLVVCVCSAVSHRNQOTWPEGIFLSSMCPINVSASTLYGIMFDAGSTGT	60
Qy	61	RIHYVTFVQKMPGOLPILGEVFDVSVKPGLSAFVDPKQGAETVQGLLEVAKDSIPRSHW	120
Db	61	RIHYVTFVQKMPGOLPILGEVFDVSVKPGLSAFVDPKQGAETVQGLLEVAKDSIPRSHW	120
Qy	121	KKTPVVLKATAGLRLPEHKAALLFEVKEIFRKSPLVPKGSVIMDGSDEGILAWTV	180
Db	121	KKTPVVLKATAGLRLPEHKAALLFEVKEIFRKSPLVPKGSVIMDGSDEGILAWTV	180
Qy	181	NFLTGQLHGRQETVGTDLGGASTQITFLPQPEKTLQTPRGYLTSEFNFNSTYKLYTH	240
Db	181	NFLTGQLHGRQETVGTDLGGASTQITFLPQPEKTLQTPRGYLTSEFNFNSTYKLYTH	240
Qy	241	SYLGFGLKAARLATLGALETGHTGHTFRSACLPRWLEAEWIFGGVYQYGGNQEVEGF	300
Db	241	SYLGFGLKAARLATLGALETGHTGHTFRSACLPRWLEAEWIFGGVYQYGGNQEVEGF	300
Qy	301	EPCVAEVLVRVGRKHOPEVQVGSFYAFSYVDRAVDTMDIDYKGGILKVEDFERKAR	360
Db	301	EPCVAEVLVRVGRKHOPEVQVGSFYAFSYVDRAVDTMDIDYKGGILKVEDFERKAR	360
Qy	361	EVCDNLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQ 400	
Db	361	EVCDNLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQ 400	

Search completed: January 29, 2004, 10:06:57  
 Job time : 16 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 09:56:59 ; Search time 34 Seconds  
(without alignments)

3248.427 Million cell updates/sec

Title: US-10-091-085-3

Perfect score: 2250

Sequence: 1 MATSWGTVPFMLVSVCSA.....ETGWALGATFHLLQSLGISH 428

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database:

SPTREMBL\_23.\*

1: sp\_archaea.\*

2: sp\_bacteria.\*

3: sp\_fungi.\*

4: sp\_human.\*

5: sp\_invertebrate.\*

6: sp\_mammal.\*

7: sp\_mhc.\*

8: sp\_organelle.\*

9: sp\_phase.\*

10: sp\_plant.\*

11: sp\_rodent.\*

12: sp\_virus.\*

13: sp\_vertebrate.\*

14: sp\_unclassified.\*

15: sp\_rvirus.\*

16: sp\_bacteriaph.\*

17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	2236	99.4	428	Q96RX0	Q96rx0' homo sapien
2	2104	93.5	407	Q8WUB3	Q8wub3' homo sapien
3	1990.5	88.5	427	Q8CD29	Q8cd29' mus musculus
4	1984.5	88.2	427	Q8BR23	Q8br23' mus musculus
5	996	44.3	483	Q8TAS7	Q8tas7' homo sapien
6	996	44.3	503	Q8N3H3	Q8n3h3' homo sapien
7	696.5	31.0	461	Q76268	Q76268' drosophila
8	696.5	31.0	464	Q9VQ18	Q9vq18' drosophila
9	616.5	27.4	479	Q9XU84	Q9xu84' caenorhabdi
10	593	26.4	278	Q8CHZ3	Q8chz3' mus musculus
11	538.5	23.9	556	Q9UT35	Q9ut35' echizoeacch
12	518.5	23.0	489	Q8H7L6	Q8h7l6' oryza sativ
13	515.5	22.9	599	Q8TGH6	Q8tgh6' candida alb
14	511	22.7	522	Q9HEM6	Q9hem6' kluyveromy
15	507.5	22.6	455	Q9SPM6	Q9spm6' medicago sa
16	503.5	22.4	467	Q9SPM7	Q9spm7' dolichos bi

#### ALIGNMENTS

##### RESULT 1

ID	Q96RX0	PRELIMINARY;	PRT;	428 AA.
AC	Q96RX0;			
DT	01-DEC-2001 (TrEMBLrel. 19, Created)			
DT	01-DEC-2001 (TrEMBLrel. 19, Last sequence update)			
DT	01-MAR-2002 (TrEMBLrel. 20, Last annotation update)			
DE	FcpH proto-oncogene protein.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=20173601; PubMed=10708485;			
RA	Recio J.A., Zambrano N., Pena, Ld, Reig J.A., Rhoads A., Rouzaut A.,			
RA	Notario V.,			
RT	"The human PCPH proto-oncogene: cDNA identification, primary			
RT	structure, chromosomal mapping, and expression in normal and tumor			
RT	cells."			
RL	Mol. Carcinog. 27:229-236 (2000).			
DR	EMBL; AF136572; AAK82950.1; .			
DR	InterPro: IPR000407; GDAI CD39_NTPase.			
DR	Pfam: PF01150; GDAI CD39_1.			
SQ	SEQUENCE 428 AA; 47431 MW; F2C4F7DE650A44F6 CRC64;			

Query Match 99.4%; Score 2236; DB 4; Length 428;  
Best Local Similarity 99.5%; Pred. No. 3.4e-185;  
Matches 426; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY	1	MATSWGTVPFMLVSVCSAVSHRNQOTWPEGIFLSSMCPINVSASTLYGIMFDAGSTGT	60
DB	1	MATSWGTVPFMLVSVCSAVSHRNQOTWPEGIFLSSMCPINVSASTLYGIMFDAGSTGT	60
QY	61	RHHVTFVQKPGQLPILGEVFDVSKPGLSAFVDQPKGAETVQGLLELVAKDSIPRSHW	120
DB	61	RHHVTFVQKPGQLPILGEVFDVSKPGLSAFVDQPKGAETVQGLLELVAKDSIPRSHW	120
QY	121	KKTPVVLKATAGLRLPEHKAKALLFEVKEIFRKSPFLVPKGSVIMDSDEGILAWTV	180

Q9xfc9 dolichos bi  
Q9spm8 lotus japon  
Q9feae pium sativ  
Q81704 arabidopsia  
Q8fvc3 glycine eo  
Q9spm5 arabidopsia  
Q8gbl1 pium sativ  
Q9m7b3 arabidopsia  
Q9agq2 arabidopsia  
Q9alv4 pium sativ  
Q8fvc2 glycine eo  
Q8rvt6 pium sativ  
Q9avm8 pium sativ  
Q8rvu0 pium sativ  
Q8fuul pium sativ  
Q8tgg8 aspergillus  
Q9aul5 medicago tr  
Q8rvt9 pium sativ  
Q8rvt8 pium sativ  
Q9c2m0 neurospora  
Q8ipz6 drosophila  
Q94ez2 arabidopsia  
Q80612 arabidopsia  
Q8cdh6 mus musculu  
Q9xi62 arabidopsia  
Q90x66 gallus gall  
Q9bhv5 leighmania  
Q8cdv7 mus musculu  
Q921q6 mus musculu

```
Db 121 KKTTPVVLKATAGLRLPEHAKALLFEVKEIFRKSPLVPKGSVIMDGSDEGLAWVT 180
Qy 181 NFLTGQHLHGRQETVGTDLGGASTQITFLPQPEKTLQTPRGYLTSEFNFSTYKLYTH 240
Db 181 NFLTGQHLHGRQETVGTDLGGASTQITFLPQPEKTLQTPRGYLTSEFNFSTYKLYTH 240
Qy 241 SYLGFGGLKAARLATLGALETGTGHTFRSACLPRWLEAWIFGGVKYQYGGNOGEVGF 300
Db 241 SYLGFGGLKAARLATLGALETGTGHTFRSACLPRWLEAWIFGGVKYQYGGNOGEVGF 300
Qy 301 EPCYAEVLVRVGRKLGHPQEVQSGSFYAFSYYYDRAVDTMDIDYKGGILKVEDFERKAR 360
Db 301 EPCYAEVLVRVGRKLGHPQEVQSGSFYAFSYYYDRAVDTMDIDYKGGILKVEDFERKAR 360
Qy 361 EVCNDLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQ 420
Db 361 EVCNDLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQ 420
Qy 421 LQSLGISH 428
Db 421 LQSLGISH 428

RESULT 2
Q8WBUB3 PRELIMINARY; PRT; 407 AA.
ID AC Q8WBUB3
DT 01-MAR-2002 (Tremblrel. 20, Created)
DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)
DE 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Ectonucleoside triphosphate diphosphohydrolase 5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Colon;
RA Strausberg R.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC020966; AAH20966.1; -.
DR InterPro; IPR000407; GDAL_CD39_NTPase.
DR Pfam; PF01150; GDAL_CD39; 1.
KW Hydrolase.
SQ SEQUENCE 407 AA; 45336 MW; D92A5F7DC9EC9E5B CRC64;

Query Match 93.5%; Score 2104; DB 4; Length 407;
Best Local Similarity 100.0%; Pred. No. 8.7e-174;
Matches 400; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MATSWGTVPFVLMVSVCSAVSHRNQQTWFEGLFLLSMCPINVSASTLYGIMFDAGSTGT 60
Db 1 MATSWGIVFVLMVSVCSAVSHRNQQTWFEGLFLLSMCPINVSASTLYGIMFDAGSTGT 60
Qy 61 RIHYVTFVQKMPGQLPILEGEVDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
Db 61 RIHYVTFVQKMPGQLPILEGEVDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
Qy 121 KKTTPVVLKATAGLRLPEHAKALLFEVKEIFRKSPLVPKGSVIMDGSDEGLAWVT 180
Db 121 KKTTPVVLKATAGLRLPEHAKALLFEVKEIFRKSPLVPKGSVIMDGSDEGLAWVT 180
Qy 181 NFLTGQHLHGRQETVGTDLGGASTQITFLPQPEKTLQTPRGYLTSEFNFSTYKLYTH 240
Db 181 NFLTGQHLHGRQETVGTDLGGASTQITFLPQPEKTLQTPRGYLTSEFNFSTYKLYTH 240
Qy 241 SYLGFGGLKAARLATLGALETGTGHTFRSACLPRWLEAWIFGGVKYQYGGNOGEVGF 300
Db 241 SYLGFGGLKAARLATLGALETGTGHTFRSACLPRWLEAWIFGGVKYQYGGNOGEVGF 300
Qy 301 EPCYAEVLVRVGRKLGHPQEVQSGSFYAFSYYYDRAVDTMDIDYKGGILKVEDFERKAR 360
Db 301 EPCYAEVLVRVGRKLGHPQEVQSGSFYAFSYYYDRAVDTMDIDYKGGILKVEDFERKAR 360
Qy 421 LQSLGISH 428
Db 421 LQSLGISH 428

RESULT 3
Q8CD29 PRELIMINARY; PRT; 427 AA.
ID AC Q8CD29
DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DE 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Ectonucleoside triphosphate diphosphohydrolase 5.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RL 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK031581; BAC27461.1; -.
SQ SEQUENCE 427 AA; 47101 MW; 6E3773C842B58477 CRC64;

Query Match 88.5%; Score 1990.5; DB 11; Length 427;
Best Local Similarity 88.1%; Pred. No. 6.5e-164;
Matches 376; Conservative 25; Mismatches 25; Indels 1; Gaps 1;

Qy 1 MATSWGTVPFVLMVSVCSAVSHRNQQTWFEGLFLLSMCPINVSASTLYGIMFDAGSTGT 60
Db 1 MATSWGAV-FMLIIACVGSFVYREQQTWFEGLFLLSMCPINVSAGTFYIGIMFDAGSTGT 59
Qy 61 RIHYVTFVQKMPGQLPILEGEVDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
Db 60 RIHYVTFVQKTAGQLPFLGEIIFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 119
Qy 121 KKTTPVVLKATAGLRLPEHAKALLFEVKEIFRKSPLVPKGSVIMDGSDEGLAWVT 180
Db 120 ERTFVVLKATAGLRLPEHAKALLFEVKEIFRKSPLVPKGSVIMDGSDEGLAWVT 179
Qy 181 NFLTGQHLHGRQETVGTDLGGASTQITFLPQPEKTLQTPRGYLTSEFNFSTYKLYTH 240
Db 180 NFLTGQHLHGRQETVGTDLGGASTQITFLPQPEKTLQTPRGYLTSEFNFSTYKLYTH 239
Qy 241 SYLGFGGLKAARLATLGALETGTGHTFRSACLPRWLEAWIFGGVKYQYGGNOGEVGF 300
Db 240 SYLGFGGLKAARLATLGALEAKGDTGHTFRSACLPRWLEAWIFGGVKYQYGGNOGEVGF 299
Qy 301 EPCYAEVLVRVGRKLGHPQEVQSGSFYAFSYYYDRAVDTMDIDYKGGILKVEDFERKAR 360
Db 300 EPCYAEVLVRVGRKLGHPQEVQSGSFYAFSYYYDRAADTHLIDYKGGVLKVEDFERKAR 359
Qy 361 EVCNDLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQ 420
Db 360 EVCNDLGSFSSGSPFLCMLDITYITALLKDGFGFADGTLQLTKKVNNIETGALGATFHL 419
Qy 421 LQSLGISH 427
Db 420 LQSLGISH 426

RESULT 4
Q8BR23 PRELIMINARY; PRT; 427 AA.
ID AC Q8BR23
DT 01-MAR-2003 (Tremblrel. 23, Created)
```

```
Db 301 EPCYAEVLVRVGRKLGHPQEVQSGSFYAFSYYYDRAVDTMDIDYKGGILKVEDFERKAR 360
Qy 361 EVCNDLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQ 400
Db 361 EVCNDLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQ 400

RESULT 3
Q8CD29 PRELIMINARY; PRT; 427 AA.
ID AC Q8CD29
DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DE 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Ectonucleoside triphosphate diphosphohydrolase 5.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RL 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK031581; BAC27461.1; -.
SQ SEQUENCE 427 AA; 47101 MW; 6E3773C842B58477 CRC64;

Query Match 88.5%; Score 1990.5; DB 11; Length 427;
Best Local Similarity 88.1%; Pred. No. 6.5e-164;
Matches 376; Conservative 25; Mismatches 25; Indels 1; Gaps 1;

Qy 1 MATSWGTVPFVLMVSVCSAVSHRNQQTWFEGLFLLSMCPINVSASTLYGIMFDAGSTGT 60
Db 1 MATSWGAV-FMLIIACVGSFVYREQQTWFEGLFLLSMCPINVSAGTFYIGIMFDAGSTGT 59
Qy 61 RIHYVTFVQKMPGQLPILEGEVDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
Db 60 RIHYVTFVQKTAGQLPFLGEIIFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 119
Qy 121 KKTTPVVLKATAGLRLPEHAKALLFEVKEIFRKSPLVPKGSVIMDGSDEGLAWVT 180
Db 120 ERTFVVLKATAGLRLPEHAKALLFEVKEIFRKSPLVPKGSVIMDGSDEGLAWVT 179
Qy 181 NFLTGQHLHGRQETVGTDLGGASTQITFLPQPEKTLQTPRGYLTSEFNFSTYKLYTH 240
Db 180 NFLTGQHLHGRQETVGTDLGGASTQITFLPQPEKTLQTPRGYLTSEFNFSTYKLYTH 239
Qy 241 SYLGFGGLKAARLATLGALETGTGHTFRSACLPRWLEAWIFGGVKYQYGGNOGEVGF 300
Db 240 SYLGFGGLKAARLATLGALEAKGDTGHTFRSACLPRWLEAWIFGGVKYQYGGNOGEVGF 299
Qy 301 EPCYAEVLVRVGRKLGHPQEVQSGSFYAFSYYYDRAVDTMDIDYKGGILKVEDFERKAR 360
Db 300 EPCYAEVLVRVGRKLGHPQEVQSGSFYAFSYYYDRAADTHLIDYKGGVLKVEDFERKAR 359
Qy 361 EVCNDLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQ 420
Db 360 EVCNDLGSFSSGSPFLCMLDITYITALLKDGFGFADGTLQLTKKVNNIETGALGATFHL 419
Qy 421 LQSLGISH 427
Db 420 LQSLGISH 426

RESULT 4
Q8BR23 PRELIMINARY; PRT; 427 AA.
ID AC Q8BR23
DT 01-MAR-2003 (Tremblrel. 23, Created)
```

```
DT 01-MAR-2003 (TEMBLrel. 23, Last sequence update)
DE 01-MAR-2003 (TEMBLrel. 23, Last annotation update)
DE Ectonucleoside triphosphate diphosphohydrolase 5.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Brain;
RX MEDLINE=22354683; PubMed=12466851;
RA The RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
DR EMBL; AK045828; BAC32507.1; -. 87BF2CC1CC1FCB9 CRC64;
SQ SEQUENCE 427 AA; 47123 MW; 47123 MW; 87BF2CC1CC1FCB9 CRC64;

Query Match 88.2%; Score 1984.5; DB 11; Length 427;
Best Local Similarity 87.8%; Pred. No. 2.2e-163;
Matches 375; Conservative 24; Mismatches 27; Indels 1; Gaps 1;

QY 1 MATSMGTVPFMLVVCVCSNVSHRNOQTHPEGIFLSMCPINVSASTLYGIMFDAGSTGT 60
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1 MATSGAV-FMLIIACVGVSTVFYREQQTFWEGVFLSSMCPINVSAGTFYGIMFDAGSTGT 59
QY 61 RIHVYTFVQKMPQQLILEGEVDSVKPGLSAFVDQPKQAEVTVQGLLEVAKDSIPRSHW 120
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
60 RIHVYTFVQKTAGQLPFLGEIIFDSVKPGLSAFVDQPKQAEVTVQGLLEVAKDSIPRSHW 119
QY 121 KKTVPVVKATAGLRLPEHKAKALLFEVKEIPRKSPLVPKGSVIMDSGDEGILAVTV 180
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
120 ERTPXLLKATAGLRLPEKQAALLLEVEIIFKNSPFLVPDGSVIMDSGYEGILAVTV 179
QY 181 NFLTGQHGHRQBTGTDLGGASTQITFLPQEKLEQTPTPGYLSFEMFNSTYKLYTH 240
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
180 NFLTGQHGHRQBTGTDLGGASTQITFLPQEKLEQTPTPGYLSFEMFNSTYKLYTH 239
QY 241 SYLGFGLKAARLATLGALETGDTGHTFRSACLPRLWEAEMIFGVKQYQGGNQEVEVG 300
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
240 SYLGFGLKAARLATLGALETGDTGHTFRSACLPRLWEAEMIFGVKQYQGGNQEVEVG 299
QY 301 EPCYAEVLVRVQKLGHPQEVORGSFYAFSYVDRAVDTDMIDYKGGILKVEDFERKAR 360
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
300 EPCYAEVLVRVQKLGHPQEVORGSFYAFSYVDRAADTHLIDYKGGVLKVEDFERKAR 359
QY 361 EVCNLENFTSGSPFLCMDSYITALLKQGFADSTVLQTKKVNNIETGVALGATFHL 420
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
360 EVCNLTGSSGSPFLCMDSYITALLKQGFADGTLQTKKVNNIETGVALGATFHL 419
QY 421 LQSLGIGS 427
DB |||||:
420 LQSLGIGT 426

RESULT 5
Q8TAS7 PRELIMINARY; PRT; 483 AA.
ID Q8TAS7
AC Q8TAS7;
DT 01-JUN-2002 (TEMBLrel. 21, Created)
DT 01-JUN-2002 (TEMBLrel. 21, Last sequence update)
DT 01-OCT-2002 (TEMBLrel. 22, Last annotation update)
DE Ectonucleoside triphosphate diphosphohydrolase 6 (Putative
DE function).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RA Strausberg R.;

Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC025980; AAH25980.1; -.
DR InterPro; IPR000407; GDA1_CD39_NTPase.
DR Pfam; PF01150; GDA1_CD39; 1.
KW Hydrolase.
SQ SEQUENCE 483 AA; 53119 MW; A8B0E5035BCDC58F CRC64;

Query Match 44.3%; Score 996; DB 4; Length 483;
Best Local Similarity 52.2%; Pred. No. 1.3e-77;
Matches 203; Conservative 57; Mismatches 123; Indels 6; Gaps 4;

QY 40 PINVSA---STLYGIMFDAGSTGTRIHYVTFVQKMPQQLILEGEVDSVKPGLSAFVDQ 96
DB ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
89 PLGTAADGHEVFGIMFDAGSTGTRVHVFOFT-RPPRETPTLTTHETFKALKPGLSAYADD 147
QY 97 PKQGAETVOGLLEVAKDSIPRSHWKTTPVVLKATAGLRLPEHKAKALLFEVKEIPRKS 156
DB ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
148 VEKSAGQIRELLDVAQDIPFPFWKATPLVLAATAGLRLPGEKAKLQKVEVKASP 207
QY 157 FLVPKGSVIMDSGDEGILAVTVNPLTQGLHGHROETVGTDLGGASTQITFLPQEK 216
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
208 FLVGDDCVSIMNGTDEGSAWITINFLTGSLKTPGSSVGMGLDLGGGSGTQIAFLPRVEGT 267
QY 217 LQTPRGYLTSEMFNSTYKLYTHSYLGLKAARLATLGALETE-GTDGHTFRSACLPR 275
DB ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
268 LQASPPGYLTALRMFNRTYKLYSYLGLGLMSARLAILGGVEGQPAKQKGLVSPCLSP 327
QY 276 WLEAEHIFGVKQYQGGNQEVEGVEPCYAEVLVRVQKLGHPQEVORGSFYAFSYYYDR 335
DB ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
328 SPKGEWEHAEVTVRVSQGAASLHELCAARVSEVLQNRVHRTVEEVKVDVFIAYSYIDL 387
QY 336 AVDTMDIDYKGGILKVEDFERKAREVCNLENFTSGSPFLCMDSYITALLKQDGFAD 395
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
388 AAGVGLIDAEGSLVVGDFEIAKVCRTLETQPOSSPSCMDLYVSLLOE-FGFPR 446
QY 396 STVLQTKKVNNIETGVALGATFHLQSL 424
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
447 SKVLKLTTRKIDNVETSWALGAIFHYIDSL 475

RESULT 6
Q8N3H3 PRELIMINARY; PRT; 503 AA.
ID Q8N3H3
AC Q8N3H3;
DT 01-OCT-2002 (TEMBLrel. 22, Created)
DT 01-OCT-2002 (TEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TEMBLrel. 23, Last annotation update)
DE Hypothetical protein (Fragment).
GN DKFZP761J1915.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Amygdala;
RA Ansgorge W., Winkner U., Mewes H.W., Weil B., Wiemann S.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL834158; CAD39864.1; -.
DR InterPro; IPR000407; GDA1_CD39_NTPase.
DR Pfam; PF01150; GDA1_CD39; 2.
KW Hypothetical protein.
FT NON TER 1
SQ SEQUENCE 503 AA; 54763 MW; 1639333F9139D8D9F CRC64;

Query Match 44.3%; Score 996; DB 4; Length 503;
Best Local Similarity 52.2%; Pred. No. 1.4e-77;
Matches 203; Conservative 57; Mismatches 123; Indels 6; Gaps 4;

QY 40 PINVSA---STLYGIMFDAGSTGTRIHYVTFVQKMPQQLILEGEVDSVKPGLSAFVDQ 96
DB ||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
109 PLGTAADGHEVFGIMFDAGSTGTRVHVFOFT-RPPRETPTLTTHETFKALKPGLSAYADD 167
```

Qy	97	PKGAEVTGVLLEVAKDSIPRSHKKTTPVVLKATAGLRLPPEHAKAKALLFEVKEIPRKSP	155
Db	168	VEKSAQCIRELLDVAQDIPDFDVFQKATPLVLKATAGLRLPPEHAKAKALLFEVKEIPRKSP	227
Qy	157	FLVPKGSVIMDGSDEGILLAVTNVFLGQLGHGRQETVGTGLDGGASTOITLPOPEKT	216
Db	228	FLVGDDCVSIMGTDEGVSAMITINFLTGSUKTPGSGVGMLDUGGSTQIAELPRVEGT	287
Qy	217	LEQTPRGYLTSPEMFNSTKYLTHSYLGFGLKAARLATLGALETE-CTDQHTFERSACLPR	275
Db	288	LQASPPGYLTALRMENRTYKLSYSLVGLGMSARLAILGGVEGQPAKDGKELVSLCP	347
Qy	276	WLEAEWIFGVKQYQVGNQGEVGFPCYAEVLRVVRGKLHQPVEVQGSFYAFSYYYDR	335
Db	348	SFKGEWAEHAEVTRYVSGOKAAASLHELCAARVSEVLQNRVHRTBEVKHVDYAFSYYYDL	407
Qy	336	AVDTDMIDYKGGILKVEDFERKAREVCNLENFTSGSPFLCNDLSITALLKDXGFGPAD	395
Db	408	AGVGLIDAEGKGSVLVGDFAEAAKYVCRTLETQPSQSPFSCMDLTVYVSLLLQE-FGFPR	466
Qy	396	STVLQLTKKVNNIETGALGATFHLLOSL	424
Db	467	SKVLKLTAKIDNVETSWALGAIPHVDSL	495
RESULT 7			
Qy	076268	PRELIMINARY; PRT; 461 AA.	
Id	076268		
Ac	076268		
Dt	01-NOV-1998 (TEMBLrel. 08, Created)		
Dt	01-NOV-1998 (TEMBLrel. 08, Last sequence update)		
Dt	01-DEC-2001 (TEMBLrel. 19, Last annotation update)		
De	NTPASE protein (LD11641P).		
Gn	NTPASE OR CG3059.		
Os	Drosophila melanogaster (Fruit fly).		
Oc	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;		
Oc	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;		
Oc	Ephydroidea; Drosophilidae; Drosophila.		
Ox	NCBI_TaxID=7227;		
Rn	[1]		
Rc	SEQUENCE FROM N.A.		
Rc	STRAIN=BERKELEY;		
Rx	MEDLINE=20196006; PubMed=10731132;		
Rx	Adams M.D., Colniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,		
Ra	Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,		
Ra	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,		
Ra	Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,		
Ra	Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,		
Ra	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,		
Ra	Abrial J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,		
Ra	Ballew R.M., Basu A., Baxendale J., Bayraktarglu L., Beasley E.M.,		
Ra	Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,		
Ra	Borkova D., Botchan M.R., Bouck J., Brockstein P., Brotter P.,		
Ra	Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,		
Ra	Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,		
Ra	de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,		
Ra	Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,		
Ra	Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,		
Ra	Rosler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,		
Ra	Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,		
Ra	Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,		
Ra	Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,		
Ra	Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,		
Ra	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,		
Ra	Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,		
Ra	Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,		
Ra	Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,		
Ra	Mout S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,		
Ra	Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,		
Ra	Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,		
Ra	Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,		
Ra	Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,		
Ra	Spier E.C., Spradling A.C., Stapleton M., Strong R., Sun E.,		



Db 295 KP-WNINGSSHSPECVGTITKNVSESEIMHLRELKSGPVLPSPYFFDRALNSGLVKGN 353

Qy 346 KGGILKVEDFERKAREVC---DNLENFTSGPFLCMDLSYITALLKDGFGFADSTVLQL 401

Db 354 EGGKIELRQFKEAEIACREKTEIDDGSHWMPQCCLDTIYISLLRDGQVQFEDNQPLVL 413

Qy 402 TKYNNIETGHALGATP 418

Db 414 AKIKGMEVSGQGLAF 430

RESULT 10

Q8CHZ3 PRELIMINARY; PRT; 278 AA.

AC Q8CHZ3

DT 01-MAR-2003 (Tremblrel. 23, Created)

DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)

DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)

DE Similar to ectonucleoside triphosphate diphosphonhydrolase 6.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

OX NCBI\_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Mammary gland;

RA Strausberg R.;

RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC038126; AAH38126.1; -

KW Hydrolase.

SQ SEQUENCE 278 AA; 30130 MW; 75A92DDIAC76297F CRC64;

Query Match 26.4%; Score 593; DB 11; Length 278;

Best Local Similarity 58.0%; Pred. No. 4.9e-43;

Matches 119; Conservative 31; Mismatches 49; Indels 6; Gaps 2;

Qy 49 YGIMFDAGSTGTRIHTVTFVQKMPGQLPILGEVDFSVKGLSAFVDPQKGAETVOGLL 108

Db 74 YGIMFDAGSTGTRIHTVTFVQKMPGQLPILGEVDFSVKGLSAFVDPQKGAETVOGLL 132

Qy 109 EVAKDSIPRSHWKTTPVLKATAGLRLPEHAKALLFEVKEIFRKS-PFLVPKGSVIMD 168

Db 133 NVAKQHPIYPDFWKATPLVLKATAGLRLPEHAKALLFEVKEIFRKS-PFLVPKGSVIMD 192

Qy 169 GSDEGILAWTVNFLTQGLHGRQETVGTLDLGASTQITFLPQFETKLEOTPRGYLTSF 228

Db 193 GTDEGVANITVNFSLTQSLKTPGSSVGMDDLGGSTQITFLPVEGTQASPPGHLTAL 252

Qy 229 EMFNSTYKLYTHSYLGFGLKAARLA 253

Db 253 QMFNRTYKLYSYRWV-----CSRLA 272

RESULT 11

Q9UT35 PRELIMINARY; PRT; 556 AA.

AC Q9UT35

DT 01-MAY-2000 (Tremblrel. 13, Created)

DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)

DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)

DE Putative guanosine-diphosphatase (Guanosine diphosphatase).

GN SPAC824.08 OR GDP1.

OS Schizosaccharomyces pombe (Fission yeast).

OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;

OC Schizosaccharomycetales; Schizosaccharomycetaceae;

OC Schizosaccharomycetes.

OX NCBI\_TaxID=4896;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=972h-;

RA Barrell B.G., Rajandream M.A., Quail M., Seegar K., Harris D.;

RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RA Sanchez R., Franco A., Notario V., Gacto M., Cansado J.;

RT "Characterization of a guanosine diphosphatase gene from Schizosaccharomyces pombe";

RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AL121741; CB57338.1; -

DR EMBL; AF465240; AAC69974.1; -

DR GeneDB SPombe; SPAC824.08; -

DR InterPro; IPR000407; GDAL CD39\_NTPase.

DR Pfam; PF01150; GDAL CD39; 1.

DR PROSITE; PS01238; GDAL CD39\_NTPASE; 1.

SQ SEQUENCE 556 AA; 61588 MW; 1D811E3D6AC6BB85 CRC64;

Query Match 23.9%; Score 538.5; DB 3; Length 556;

Best Local Similarity 34.4%; Pred. No. 7.7e-38;

Matches 144; Conservative 61; Mismatches 157; Indels 57; Gaps 13;

Qy 49 YGIMFDAGSTGTRIHTVTFVQKMPGQLPILGEVDFSVKGLSAFVDPQKGAETVOGLL 108

Db 134 YVLMIDAGSTGSRVHVYQFNCPNPS--PKLEEFFKMIIEFGLSFAGDPEGAASLDPLL 191

Qy 109 EVAKDSIPRSHWKTTPVLKATAGLRLPEHAKALLFEVKEIFRKS-PFLVPKGSVIM 167

Db 192 DYAMENVPPEYRRCSPIAVKATAGLRLTGESEAKALKSVRQHLNDYPPPIVKGVISIL 251

Qy 168 DGSDEGILAWTVNFLTQGLHGRQETVGTLDLGASTQITFLPQFETKLEOTPRG--- 223

Db 252 EGSMEGIYAWITINYLGLTGKATHTSTVAVMDLGGASTQLVPEPRFASDGESLVDGDHK 311

Qy 224 YLTSFEMFNSTYKLYTHSYLGFGLKAARL-----ATLGALETGTDGHTFRSAC 272

Db 312 YVLDYN--GEQYELYQSHLYGLKEARKLIHKFVLNNAEALKESLELLG-DSTSIHPC 368

Qy 273 L-----PWLAEWIFGVGYQYGNQSGEVGFEPCEYAEVLRVVR---GKLHQP 318

Db 369 LHUNASLTHPDSKSEASEVVFVGPSLAHLSQLCRGIAEKALYKDKNCVPCSFVGVQHP 428

Qy 319 ---EYVQSGFYAFSYDYDRAVDTMDIYKGGILKVEDFERKAREVC----- 363

Db 429 KFTETFTDPSIYLSYFYDR----MISLGMPTFTIEDMKYLSVCSGPTYWQDAFSL 483

Qy 364 -DNLENFTSGPFLCMDLSYITALLKDGFGFADSTVLQTKKNNIETGHALGATFHL 421

Db 484 TDALKELKE-EPEWCLDLNMISSLVGYEIPNNRLHTAKKIDNKLGLWCLGASLSML 541

RESULT 12

Q8H7L6 PRELIMINARY; PRT; 489 AA.

AC Q8H7L6

DT 01-MAR-2003 (Tremblrel. 23, Created)

DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)

DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)

DE Putative apyrase.

GN OSJNB0014110.10.

OS Oryza sativa (Japonica cultivar-group).

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

OC Ehrhartoideae; Oryzaceae; Oryza.

OX NCBI\_TaxID=39947;

RN [1]

RP SEQUENCE FROM N.A.

RA Wing R.A., Yu Y., Soderlund C., Kim H.-R., Rambo T., Saski C.;

RA Currie J., Collura K.;

RT "Rice Genomic Sequence."

RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; AC126222; AAN65004.1; -

SQ SEQUENCE 489 AA; 52799 MW; 0BF6BF154488D38E CRC64;

Query Match 23.0%; Score 518.5; DB 10; Length 489;

Best Local Similarity 34.0%; Pred. No. 3.4e-36;

Matches 143; Conservative 70; Mismatches 149; Indels 59; Gaps 16;







GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 09:58:54 ; Search time 16.5 Seconds  
(without alignments)  
2494.557 Million cell updates/sec

Title: US-10-091-085-3  
Perfect score: 2250  
Sequence: 1 MATSWGTVFPLVWVCVCSA.....ETGWALGATPHLLQSLGISH 428

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 76.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	616.5	27.4	479	2 T23508	hypothetical prote
2	538.5	23.9	556	2 T39109	probable guanidine
3	489	21.7	455	2 S48859	nucleoside triphos
4	483.5	21.5	518	2 A40732	guanosine-diphosph
5	446.5	19.8	454	2 JC4616	aprase (EC 3.6.1.
6	425.5	18.9	516	2 G84442	probable nucleosid
7	397	17.6	483	2 D86276	hypothetical prote
8	387.5	17.2	485	2 T31477	hypothetical prote
9	372.5	16.6	557	2 T16696	hypothetical prote
10	372	16.5	630	2 S50463	hypothetical prote
11	351	15.6	510	2 T56242	lymphoid cell acti
12	347	15.4	572	2 T40856	probable nucleotid
13	329.5	14.6	405	2 E86276	hypothetical prote
14	274	12.2	1052	2 T04439	hypothetical prote
15	248	11.0	508	2 C86276	7A19.33 protein -
16	138	6.1	628	2 A55421	nucleoside-triphos
17	110	4.9	497	1 JC2192	subtilisin-like pr
18	106.5	4.7	553	1 G1BP5V	gene 1 protein - s
19	105.5	4.7	774	2 T14555	DNA polymerase hom
20	104.5	4.6	684	2 T47694	probable serine/thr
21	104.5	4.6	4427	2 PN0637	polyketide synthas
22	102.5	4.6	1019	2 T40813	probable cell divi
23	101.5	4.5	535	2 F97910	glucan 1,6-alpha-g
24	101.5	4.5	1151	2 H71347	hypothetical prote
25	99.5	4.4	535	2 C95040	glucan 1,6-alpha-g
26	99.5	4.4	548	1 HYBSS	bacillolysin (EC 3
27	97.5	4.3	604	2 F89453	protein F35H2.4 [
28	97.5	4.3	654	2 AG3522	iron-regulated out
29	97	4.3	1220	2 T18291	patched protein -

30	96	4.3	711	2 S66749	hypothetical prote
31	95.5	4.2	2076	2 S15999	fatty-acyl-CoA syn
32	95	4.2	652	1 JC2191	subtilisin-like pr
33	93	4.1	478	2 F90497	hypothetical prote
34	93	4.1	635	1 T37835	probable phosphoes
35	93	4.1	962	2 JC5571	subtilisin-like pr
36	93	4.1	969	1 A39490	subtilisin-like pr
37	93	4.1	975	2 JC5570	coproporphyrinogen
38	92	4.1	290	2 A87312	H+-transporting tw
39	92	4.1	528	2 S13641	glycine hydroxymet
40	91	4.0	417	2 H70174	t-plasminogen acti
41	91	4.0	477	2 JS0597	gluconate kinase g
42	91	4.0	500	2 D83984	gluconate kinase g
43	91	4.0	726	2 AB0122	probable ferric ei
44	90.5	4.0	301	2 E71482	phosphatidylserine
45	90.5	4.0	377	2 C82113	succinyl-diaminopi

ALIGNMENTS

RESULT 1

T23508  
hypothetical protein K08H10.4 - Caenorhabditis elegans  
C;Species: Caenorhabditis elegans  
C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 17-Mar-2000  
C;Accession: T23508  
R;Gardner, A.  
submitted to the EMBL Data Library, November 1996  
A;Reference number: Z19750  
A;Accession: T23508  
A;Status: preliminary; translated from GB/EMBL/DDBJ  
A;Molecule type: DNA  
A;Residues: 1-479 <WIL>  
A;Cross-references: EMBL:Z83113; PIDN:CAB05544.1; GSPDB:GN00023; CESP:K08H10.4  
A;Experimental source: clone K08H10  
C;Genetics:  
A;Gene: CESP:K08H10.4  
A;Map position: 5  
A;Introns: 36/3; 83/3; 189/1; 300/2; 412/3  
C;Superfamily: nucleoside triphosphatase chromatin-associated

Query Match	27.4%	Score	616.5;	DB 2;	Length	479;
Best Local Similarity	34.3%	Pred. No.	1.3e-43;			
Matches	150;	Conservative	81;	Mismatches	167;	Indels 39; Gaps 16;
Qy	9	FPMVVS--CVCASVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGTRIHYVT	66			
Db	6	FSILLISFFSLJSVVTTKTQY-WCHGDGVLN-----NQHTCRFFTIVIDAGSTGTRLHLK	60			
Qy	67	FVQK-----MPGQLPILEGEVFDVSKPGLSAFVQPKQGAETVQGLLEVAKDSIPRSH	119			
Db	61	FHDPAIASHGMPFK---VEKEIFQEVKGLSLSPSAADSLPFLORARKEVPHFM	117			
Qy	120	WKPTPVVLKATAGRLRLPEHKAKALLFEVKE-IPRKSPPFLVPKGSVIMDGSDEGLAWV	178			
Db	118	WEXTPTLTKATAGRLRLPGMADDILESVEERIFNSGFFAAFPDPAVNVMPGSDGVYSWF	177			
Qy	179	TVNFLTGQLH-----GHR---QETVGLDLCGASTQITFLPQFEXTLEQTPRGYLTSP	229			
Db	178	TLNILELFTDPTVGHKPAHRVAAFDLGGGSGTQITVYMPNNEAVFSHVH-GYERDID	236			
Qy	230	MFNSTVKLYTHSYLGLKAARLATGALTEG--TDGHTFRSACLPRMLE-AEWIFGGV	286			
Db	237	FFGHRLRFTSHFLGNGLAARLNIL-QLETDNEISTHQLITSCMPEGQLTEWEY-AL	294			
Qy	287	KYOYGNQGEVGFBCPYAEVLVRV-RKXHQPEEVQVRSFYAFSYYYDRAVDTMDIYE	345			
Db	295	KF-WNINGSSSHFSFCYGTGTTKNFVSESSEIMHLRELKGSPLYFSYFFDRALNSGLVKGN	353			
Qy	346	KGILKVEDEPERKAREVC-----DNLENFTSGSFLCNDLSYITALLKDGFGPADSTVLQ	401			
Db	354	EGGKIELRQPKAAEIAACREKTEIDDGSHMWPQCLDLTYIYSLRDGQYQFEDNOPLVL	413			

QY 402 TKVNNIETGALGATF 418  
 Db 414 AKKIKMEVSGOGLAF 430

RESULT 2  
 T39109  
 probable guanosine-diphosphatase - fission yeast (Schizosaccharomyces pombe)  
 C:Species: Schizosaccharomyces pombe  
 C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999  
 C:Accession: T39109  
 R:Barrell, B.G.; Rajandream, M.A.; Quail, M.; Seegar, K.; Harris, D.  
 submitted to the EMBL Data Library, October 1999  
 A:Reference number: Z21828  
 A:Accession: T39109  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-556 <BAR>  
 A:Cross-references: EMBL:AL121741; PIDN:CAB57338.1; GSPDB:GN00066; SPDB:SPAC824.08  
 A:Experimental source: strain 972h-; cosmid c824  
 C:Genetics:  
 A:Gene: SPDB:SPAC824.08  
 A:Map position: 1

Query Match 23.9%; Score 538.5; DB 2; Length 556;  
 Best Local Similarity 34.4%; Pred. No. 5.7e-37;  
 Matches 144; Conservative 61; Mismatches 157; Indels 57; Gaps 13;

QY 49 YGIMFDAGSTGTRIHVYTFVQKMPGQLPILLEG-EVFDVKPGLSAFVDPQKQAGTVOGLL 108  
 Db 134 YVLMIDAGSTGSRVHVYQFNCFNS--PKLEEFKMIPLGLSSFAGDPEGAAASLDPLL 191  
 QY 109 EVAKDSIPRSHWKTTPVVLKATAGLRLLPEHKAKALLFEVKEIFPKS-PFLVPKGSYSIM 167  
 Db 192 DYAMENVPERRCSPVAVKATAGLRGSEAKALKSVKQHLNDYPPFIVKDGVSIL 251  
 QY 168 DGSDEGILAWTVNFLTQGLHGH-ROETVGTDLGGASTQITFLPQPEKTELEOTPRG--- 223  
 Db 252 EGSMEGIYAVITINVLTLGKGKATHSTVAVMDLGGASTQLVFEPFASDGSGLVDGDHK 311  
 QY 224 YLTSEFMENSTYKLYTHSYLGFGLKAARL-----ATLGALETEGDTGHTFRSAC 272  
 Db 312 YVLDYN--GEQELYQHSLGYGLKEARKLTHKEVLNNAEALKESLLG-DSTSIITPC 368  
 QY 273 L-----PRMLEAEWIFGGVKYQYGGNQEVEGFEPCYAEVLVRV---GKLLHOP 318  
 Db 369 LHLNLSLTHPDSKSEAEVWFVGSLSLAHLSLQCRGIAEKALYKDKNCVPRPCSPNGVHQ 428  
 QY 319 ---EVQSGSPYAFSYYYDRAVDTDMDIDYKGGILKVEDFERKAREVC----- 363  
 Db 429 KFTETFTDSPILSYFYDR-----MISLGMPSFTIEDMKYLSVCSGPTYWQDAFSL 483  
 QY 364 -DNLENFTSGSPFFICMLDSYITALLKDGFGFADSTVQLTKVNNIETGALGATFHLL 421  
 Db 484 TDALKELKE-BPEWCLDINYMISLSYGYEIPNNRQLHTAKKIDNKELGWLCLGASLSML 541

RESULT 3  
 S48859  
 nucleoside triphosphatase precursor, chromatin-associated - garden pea  
 C:Species: Pisum sativum (garden pea)  
 C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 29-Sep-1999  
 C:Accession: S65147; S48859  
 R:Haieh, H.L.; Tong, C.G.; Thomas, C.; Roux, S.J.  
 Plant Mol. Biol. 30, 135-147, 1996  
 A:Title: Light-modulated abundance of an mRNA encoding a calmodulin-regulated, chromatin  
 A:Reference number: S65147; MUID:96197404; PMID:8616230  
 A:Accession: S65147  
 A:Molecule type: mRNA  
 A:Residues: 1-455 <HS2>  
 A:Cross-references: EMBL:Z32743; NID:G563611; PIDN:CAA83655.1; PID:G563612  
 C:Superfamily: nucleoside triphosphatase chromatin-associated

C;Keywords: nucleus

Query Match 21.7%; Score 489; DB 2; Length 455;  
 Best Local Similarity 33.7%; Pred. No. 6e-33;  
 Matches 140; Conservative 62; Mismatches 160; Indels 54; Gaps 14;

QY 49 YGIMFDAGSTGTRIHVYTFVQKMPGQLPILLEG-EVFDVKPGLSAFVDPQKQAGTVOGL 107  
 Db 44 YAVVFDAGSTGSRHVYVFNQNL-DLLHIGKGYEYNNKITPGLSSVANNPEQAASLIPL 102  
 QY 108 LEVAKDSIPRSHWKTTPVVLKATAGLRLLPEHKAKALLFEVKEIF-RKSPFLPKGSVSI 166  
 Db 103 LEQAEADVPPDQKTPFVRGATAGLRLLNGDSEKILQSVRDMLSNRSTFNVPQDAVSI 162  
 QY 167 MDGSEGILAWTVNFLTQGLHGHROETVGTDLGGASTQITFLPQPEKTELEOTPR--- 222  
 Db 163 IDGTQEGSYLWTVNVALGNLKKYTKTVGVIDLGGSVQMAVAVS-KKTAKNAPKAVDG 221  
 QY 223 --GYLTSEFMENSTYKLYTHSYLGFGLKAARLALATLGALETEGDTGHTFRS--ACLPRWLE 278  
 Db 222 DDPYIKKVLKGIPIYDLVHSHYLFHFGREASRAEILKL-----TPRSPNCLLAGFN 272  
 QY 279 AEWIFGGVKYQYGGNQEVEGFEPCYAEVLVRVGRKLHOPPEEVQSGF----- 326  
 Db 273 GIYTGEEFKATAYTSG-ANFNKCKNTIRKAL--KLNYPCTYQNCCTFGGIWNGGNGNQ 329  
 QY 327 ---YAFSYYYDRAVDTDMDIDYKGG-ILKVEDFERKAREVCD-NLENFTSGSPFL----- 376  
 Db 330 KNLFASSSFFYLPEDTCMDVASTPNFILRPVDIETKAEKALNFEADAKSTYFFLDKKNV 389  
 QY 377 ----CMDSYITALLKDGFGFADSTVQLTKVNN-----IETGALGATFHLLQSL 424  
 Db 390 ASYVCMDLIYQYVLLVDGFLDPLQKITSGKTEYQDAIVEAAWPLGNAVEAISAL 445

RESULT 4  
 A40732  
 guanosine-diphosphatase (EC 3.6.1.42) - yeast (Saccharomyces cerevisiae)  
 N:Alternate names: protein YEL042w  
 C:Species: Saccharomyces cerevisiae  
 C:Date: 21-Sep-1993 #sequence\_revision 18-Nov-1994 #text\_change 02-Sep-2000  
 C:Accession: A40732; B40732; S30837; S50502  
 R:Abelion, C.; Yanagisawa, K.; Mandon, E.C.; Hausler, A.; Moremen, K.; Hirschberg, C.B.;  
 J. Cell Biol. 122, 307-323, 1993  
 A:Title: Guanosine diphosphatase is required for protein and sphingolipid glycosylation  
 A:Reference number: A40732; MUID:93308137; PMID:8391537  
 A:Accession: A40732  
 A:Molecule type: DNA  
 A:Residues: 1-518 <ABE>  
 A:Cross-references: EMBL:L19560; NID:G349392; PIDN:AAA34656.1; PID:G349393  
 A:Note: sequence extracted from NCBI backbone (NCBIN:134708, NCBIIP:134711)  
 A:Accession: B40732  
 A:Molecule type: protein  
 A:Residues: 125-144;238-257;276-281;366-374;399-412 <AB2>  
 R:Mulligan, J.T.; Dietrich, F.S.; Hennessey, K.M.; Sehl, P.; Komp, C.; Wei, Y.; Taylor,  
 submitted to the EMBL Data Library, February 1993  
 A:Reference number: S30812  
 A:Accession: S30837  
 A:Molecule type: DNA  
 A:Residues: 1-518 <MUL>  
 A:Cross-references: GB:U18779; EMBL:U10830; NID:G603625; PIDN:AAB65000.1; PID:G603637  
 R:Dietrich, F.S.  
 submitted to the EMBL Data Library, December 1994  
 A:Description: The sequence of S. cerevisiae cosmid 8199, 8334, and 9871.  
 A:Reference number: S50491  
 A:Accession: S50502  
 A:Molecule type: DNA  
 A:Residues: 1-518 <DIE>  
 A:Cross-references: EMBL:U18779; NID:G603625; PIDN:AAB65000.1; PID:G603637; MIPS:YEL042w  
 C:Genetics:  
 A:Gene: SGD:GDAL  
 A:Cross-references: SGD:S0000768; MIPS:YEL042w  
 A:Map position: 5L



Db 61 NLRSLRYVVVDGGSTGTTRIHFVGY--RIESGKPFVFRGANYASLKLHPLGSLSAFADDP 118  
Qy 98 KQGAETVOGLEVAKDSIPRSHWKTTPVVLKATAGLRLPEHAKALLPEVKEIFRSPF 157  
Db 119 DGASVUTELVFPKGRVPGMWJETEVRLMATAGMLLELPVQEKILGVARRVLKSSGF 178  
Qy 158 LVPKGSVIMDGSDEGILAWVTNVLFTQLGHGRQETVGTDLGGASTQITFL-----P 211  
Db 179 LFRDEMASVSGSDEGVYAWVANFALGSLGGDPLKTTGIVELGGASQVTFVSSEMP 238  
Qy 212 QFEKLTQTPRGYLTSEMFNSTYKLYTHSYLGLKAARLATLAL-----ETGTD 264  
Db 239 EFSRTI-----SFG--NVTYLSHSGFLHFGQNAADHKLWGSLLSRDHNSAVEPTR 287  
Qy 265 GHTFRSACLPR-----WLEAEWIFGVKYQYGNQGBGEVFPFCVAEVLRVV 311  
Db 288 EKIPDPCAPKGYNLDANTOKHLSGLLAESRSDSFQAGN-----YSOCRSAULTIL 341  
Qy 312 RGKLHQPVEEVQVGSFYAFSYYVYDRAVDTDMI-----DYKGGILKVEDFERKAREV 362  
Db 342 QDNGNRIILIIAGSFPLFGIGEKAWLSNLSAGERFCGEDWSK--LRVKDPSLHEEDL 398  
Qy 363 CDNLNFTSGSPFLCMLDSYITALLKDGFGP-ADSTVLQLTKKVNNIETGHWALGA 416  
Db 399 LR-----YCFSSAVIVSLHDTLGTPLDDEIRIGYANOAGDIPLDWALGA 442

RESULT 7  
D86276  
hypothetical protein F7A19.34 [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 18-Nov-2002  
C:Accession: D86276  
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
Chan, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;  
ansen, N.F.; Hughes, B.; Huizlar, L.  
Nature 408, 816-820, 2000  
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A:Authors: Salzberg, S.L.; Schwarz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A:Reference number: A86141; MUID:21016719; PMID:11130712  
A:Accession: D86276  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-483 <STO>  
A:Cross-references: GB:AB005172; NID:G5080801; PIDN:AAD39311.1; GSPDB:GN00141  
C:Genetics:  
A:Map position: 1  
C:Superfamily: nucleoside triphosphatase chromatin-associated

Query Match 17.6%; Score 397; DB 2; Length 483;  
Best Local Similarity 27.6%; Pred. No. 3.3e-25;  
Matches 120; Conservative 71; Mismatches 162; Indels 82; Gaps 17;  
Qy 32 GIFLSSMCPINVSASTLYGIMFDAGSTGTRIHV--YTFVQKMPQLPILGEVFD-----84  
Db 59 GSLSRRCKLR-----YSVLIDAGSSGTRVHVFGYWFESGKP-----VDFDGEKH 103  
Qy 85 ----SVKPKGSFAPVDQKQAGETVQGLEVAKDSIPRSHWKTTPVVLKATAGLRLPEHK 140  
Db 104 YANUKLTPGLSSYADNPEGASVSTKLVEFAKQIPKRMFRSRDIRLMATAGMLLEVPV 163  
Qy 141 AKALLFEVKEIFRKSPLVPKGSVIMDGSDEGILAWVTNVLFTQLGHGRQETVGTDL 200  
Db 164 QEILEVTRVLRSSGFMFRDEWANVISGSGDEGIYSMTANYALGSLGTDPLETTGIVEL 223  
Qy 201 GGASTQITFLDQPKETLEOTPRGYLTSEMFNSTYKLYTHSYLGLKAARLATLGALET 260  
Db 224 GGASQAVTFV-----SSEHVPPEYSRTIAYGNISYTIYSHSFLDYGKDAALKLLEKLQN 278

Qy 261 EG---TDGHTFRSACLPRWLEAEWIF--GGVKYQYG-----GNQGEVGEFCEPCYAE 306  
Db 279 SANSTVDG--VVEDPCTPK---GYIYDTSKNYSYSGFLADESKLKGSLQAGNFSKCRSA 333  
Qy 307 VLRVVR-----GKLHQPVEEVQVGSFYAFSYYVYDRAVDTDMIYKGGILKV- 352  
Db 334 TPAALLKKGKENCILYHCSTGSTFTPD--LQGSPLATASFYYTAKFEL--EEKWJSELI 389  
Qy 353 -----EDFERKAREVCNLENFTSGSPFLCMLDSYITALLKDGFGP--DSTVLQLT 402  
Db 390 PAKGRYCGEWSKLIILEYPTTDEYLRG--YCFSAAYTISMLHDSLGIALDDSEITYAS 446  
Qy 403 KK-VNNIETGHWALGA 416  
Db 447 KAGEKHIPLDWALGA 461

RESULT 8  
T34147  
hypothetical protein C33H5.14 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 17-Mar-2000  
C:Accession: T34147  
R:Bradshaw, H.; Stellyes, L.  
submitted to the EMBL Data Library, November 1995  
A:Description: The sequence of C. elegans cosmid C33H5.  
A:Reference number: Z21482  
A:Accession: T34147  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-485 <BRA>  
A:Cross-references: EMBL:U41007; PIDN:AAA82272.1; CESP:C33H5.14  
A:Experimental source: strain Bristol N2  
C:Superfamily: nucleoside triphosphatase chromatin-associated

Query Match 17.2%; Score 387.5; DB 2; Length 485;  
Best Local Similarity 27.7%; Pred. No. 2.1e-24;  
Matches 120; Conservative 84; Mismatches 152; Indels 77; Gaps 18;  
Qy 49 YGIMFDAGSTGTRIHYVTFVQKMPGQLPILGEVFD-----VKPGLSAFVDQKQAG 102  
Db 25 YGVICDAGSSGTRLFVVT-LKPLSGGLTNIDTLIHESEPWWKVTPLGSLSGDKPEQVVE 83  
Qy 103 TVQGLLEVAKDSIPRSHWKTTPVVLKATAGLRLPEHAKALLFEVKEIFRKSPLF-VPK 161  
Db 84 YLTPLRFABEHIPIYBQLGETDILLIFATAGMRLLEPAQKDAIITKLNQGLKSVTALRVSD 143  
Qy 162 GSYSIMDGSDEGILAWVTNVLFTQLGHGRQETVGTDLGGASTQITFLPQFEK-----215  
Db 144 SNRIIDGAWEGIYSWIAVNIILGRFDKENDSKVGMIDMGASVQIATFANKEVSYNGG 203  
Qy 216 TLBQTPRGYLTSEMFNSTYKLYTHSYLGLKAARLATLGALETGTDGHTFRSACLPR 275  
Db 204 NVYEINLGSITENEDYK--YKIYSTTFLGVAGNEGLKKYENSLVKSGNS-----NDSCSPR 257  
Qy 276 WLEAEWIFGVKYQYGNQGBGEVFPFCVAEVLRVVVRGKLHQPE-----319  
Db 258 GLNR--LIG-----EFTVNGTGE--WDVCLAQVSLI-GDKAQPSCNPCTFLNRVIAPSV 308  
Qy 320 EVQGSFYAFS--YYDRAVDTDMIYKGGILKVEDFERKAREVC-----NLENFTSGSP 374  
Db 309 NLSTVOLYGFSEWYTTIS-----NFGSGGEYHYQFTDEVKVKYCKQDWNIDQGFPRNE 362  
Qy 375 F-----LCMDLSYITALLKDGFGFADST--VLQLTKKVNNIETGHWALGA-----416  
Db 363 FPNADIERLGTNCFKAAWVTSVLHDGFN-VDKTKHLFQSVLKIAGEBMQWALGAMLYHSK 421  
Qy 417 --TFHLLQSLGIS 427  
Db 422 DLKFNLLQLEVA 434

RESULT 9  
T16696  
hypothetical protein R07E4.4 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C>Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Sep-1999  
C:Accession: T16696  
R:Miller, N.  
submitted to the EMBL Data Library, October 1995  
A:Description: The sequence of C. elegans cosmid R07E4.  
A:Reference number: Z18561  
A:Accession: T16696  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-557 <ML>  
A:Cross-references: EMBL:U39652; NID:g1049390; PID:g1049394; PIDN:AAA80403.1; CESP:R07E4  
A:Experimental source: strain Bristol N2  
C:Genetics:  
A:Gene: CESP:R07E4.4  
A:Introns: 39/1; 67/1; 145/2; 192/1; 244/3; 317/3; 451/3; 488/2

Query Match 16.6%; Score 372.5; DB 2; Length 557;  
Best Local Similarity 28.3%; Pred. No. 4.6e-23;  
Matches 126; Conservative 69; Mismatches 160; Indels 91; Gaps 21;

QY 49 YGIMFDAGSTGTRIHYVTFVOKMPGOLPILEGEVFDSD-----VKPGLSAFVDPQKQAE 102  
DB 44 YGVICDAGSTGTRLFVYNWISTDSSELIQIEPIYDKNPKWIKISPOLSTFGTKPAQAE 103

QY 103 TVQGLLEVAKDSIPRSHWKKTTPVVLKATAGLRLLPEH-----KAKALLFEVKEIFRKSPP 157  
DB 104 YLRPLMELAEHRHPEEKRPVTPVIFATAGMRLIPDEYVLIGQEAVAL---KNLRNKLPK 160

QY 158 L-----VPKGSVSMDSDEGLAWTVNFLTGLQLH-----GH-RQSTVGTGLD 200  
DB 161 ITSMQVLKEHRIIEGKWEIGYIAVNYALGFNKFTATLDFPGTSPAHARQKTVGMIDM 220

QY 201 GGASTQITF-LPQFEK---TLQTPRGYLTSPFEMENSTVKLYTHSVLGF-----GLKAAR 251  
DB 221 GGASQIAFELPDTDSFSSINVENINLGCRDSDLFK--YKLFVTFLGYGVNEGIRKYE 278

QY 252 LATLGALETEGTDGTFPSACLPRLWEAEWIFGVKQYQGN--QEG----- 296  
DB 279 HMLLSKLKQD--NGTVIQDDCMPLNLHKT-----VTLENGENFVRRTGNWNTCSNEVKK 331

QY 297 -----EVGFPCYAEVLVRVGKLUHQP-EVQSGFYAFSYVYDRAVDTMDIYEKGGILK 351  
DB 332 LLNPESSEYCKAEAAKCYFGAVPAPSIPLSNIEYGFSEYWTSHDV-----LGLGGQYD 387

QY 352 VEDFERKAREVCD-----NLENFTSGSP-----FLCWDLSYITALLKDGFGFADST 397  
DB 388 AENIAKTKQYCSKRWSTIQAESKKQIYPADBEERLTQCFKSAWITSVLHDGFS-VDKT 446

QY 398 --VLQLTKVNNIETGWALGA-TFHL 420  
DB 447 HNKEFSVSTTAGQEVQVQALGMIVHM 472

RESULT 10  
S50463  
hypothetical protein YER005w - Yeast (Saccharomyces cerevisiae)  
C:Species: Saccharomyces cerevisiae  
C>Date: 28-May-1993 #sequence\_revision 24-Feb-1995 #text\_change 19-Apr-2002  
C:Accession: S50463  
R:Dietrich, F.S.  
submitted to the EMBL Data Library, December 1994  
A:Description: The sequence of S. cerevisiae cosmid 9537, 9581, 9495, 9867, and lambda  
A:Reference number: S50433  
A:Accession: S50463  
A:Molecule type: DNA  
A:Residues: 1-630 <DIE>  
A:Cross-references: EMBL:U18778; NID:g603592; PID:g603597; GSPDB:GN00005; MIPS:YER005w

C:Genetics:  
A:Gene: SGD:YND1; MIPS:YER005w  
A:Cross-references: SGD:S0000807  
A:Map position: 5R

Query Match 16.5%; Score 372; DB 2; Length 630;  
Best Local Similarity 26.2%; Pred. No. 6.1e-23;  
Matches 117; Conservative 73; Mismatches 165; Indels 92; Gaps 17;

QY 49 YGIMFDAGSTGTRIHYVTF-----VQKMPGOLPILEGE---VFDSPVKPGLSA 92  
DB 10 FGIVIDAGSSGRIHVFKQDTESSLHATNQDSQSLOSVPHIHQEKDWF-KLNFGLS 68

QY 93 FVDQPKQGAET-VQGLLEVAKDSIPRSHWKKTTPVVLKATAGLRLLPEHKAALLFV-KE 150  
DB 69 FEKKPQDAYKSHIKPLLDPAKNIIPESHWSSCPVFQATAGMRLLPQDIQSSILDGLCOG 128

QY 151 IFRKSPFLVP--KGSVMSDSEGLAWTVNFLTGLQLHGRQE-----TVGTLDLGA 203  
DB 129 LKHPAEFLVEDCSAQIQVIDGETGLYGLNLNLYGHFNFDYNPEVSDHFTFGFMDMGA 188

QY 204 STQITFLPQFEKTL-----OTPRGYLTSPFEMENSTVKLYTHSVLGLFGLKAARL 252  
DB 189 STQIAPAPHSGEIAHRDDIATIFLRSVNGDLQKWDVFSV-----WLGFGANQARR 241

QY 253 ATLGA-----ETEGTDGHTFPSACLPRLWEAEWIFGVKQYQGNQGEVGP 301  
DB 242 RYLAQLINTLPENTNDYENDDFSTRNLNDCMPRGSSDTDFPKDTIFHLAGSGNYEQCTK 301

QY 302 PCYAEVL-----VVRGKLHQP-EVQSGFYAFSYVYDRAVDTMDIYEKGGILKV 352  
DB 302 SIYPLLLKNMPCDDEPCLFNGVHAPRIDFANDKFIGTSEYWTYANDV-----FKLGGYNF 357

QY 353 EDPERKAREVCD-----LENFTSG-----SPFL---CWDLSYITALLKDGFGFA--- 394  
DB 358 DRPKSLRLEFCNSNWTQILANSKGVYNSIPENFLKDACFKGNWLVNLHGEFDMRIDV 417

QY 395 -----DSTVLQTLTKVNNIETGWALG 415  
DB 418 DAENVNDRPLFQSVKVERELSWTLG 444

RESULT 11  
I56242  
lymphoid cell activation antigen - human  
C:Species: Homo sapiens (man)  
C>Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 18-Nov-2002  
C:Accession: I56242  
J: Immunol. 153, 3574-3583, 1994  
A:Title: The CD39 lymphoid cell activation antigen. Molecular cloning and structural char  
A:Reference number: I56242; MUID:95015846; PMID:7930580  
A:Accession: I56242  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-510 <RES>  
A:Cross-references: GB:S73813; NID:g765255; PIDN:AAB32152.1; PID:g765256  
C:Superfamily: nucleoside triphosphatase chromatin-associated

Query Match 15.6%; Score 351; DB 2; Length 510;  
Best Local Similarity 27.8%; Pred. No. 2.6e-21;  
Matches 123; Conservative 71; Mismatches 175; Indels 74; Gaps 19;

QY 32 GTFLLSMKPCINVSASTLYGIMPDAGSTGTRIHYVTFVQKMPGOLPILEGEVFDSPVK-PGL 90  
DB 36 GLTONKALPENVK-----YGLVDAGSSHTSLYIKRPAEKENDTGVVHVQVEECRKGPI 91

QY 91 SAFVDOPKGAETVQGLLEVAKDSIPRSHWKKTTPVVLKATAGLRLL---PEHKAALLFE 147  
DB 92 SKFVQVKEIGIVLTCMERAREVIRPSQHQETPVVLGATAGMRLLRMESEELADRLDV 151

QY 148 VKEIFRKSPLVPKGSVSMDSDEGLAWTVNFLTGLQ-----HGRQETV 195

Db 152 VERSLSNVPF--DQGARITIQBEGAYGWTINYLKFSQKTRWFSIYPYETNNQETP 209  
 Qy 196 GTDLGGASTQITFLPQFTELETPRGYTSFEMFNSTYKLYTHSYLGLFGKLAARLATL 255  
 Db 210 GALLDGGASTQITFLPQ-NTIE-SPDNAL-QFRLYGKQYNNVYTHSYFLCYGKQALWQKL 266  
 Qy 256 GALTETGDTGHTFRSACL-----PRWLEAEWIFGQVKYQYGGNQE 297  
 Db 267 -AKDIOVASNEILRDPCHPGYKKVNVVSDLYKTPCKRFEMTLFPOQFEIQGIN---- 321  
 Qy 298 VGFPEPCYAEVLRV-----VRKLUHQPEEVQSGSYAFSYYYDRAVDTMDIYE 345  
 Db 322 --YQCHOSILELNTSVCPYSCQAFNGIFLPLQDGFAPSAP-YFWMKPLN---LTSE 375  
 Qy 346 KGGILKVEDFERK-AREVCNLENFTSG--SPFL---CMDSYITALLKDGFGF-ADS-T 397  
 Db 376 KVSEKTEMMKCAQWBEIKTSYAGVEKYLSEYCFSTYILSLLLQYHHTADSW 435  
 Qy 398 VLQTKKNNIETGALGATHPL 420  
 Db 436 HIHFITGKIQSDAGWTGLGYMLNL 458

RESULT 12  
 T40856  
 probable nucleotide phosphatase - fission yeast (Schizosaccharomyces pombe)  
 C:Species: Schizosaccharomyces pombe  
 C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999  
 C:Accession: T40856  
 R:Ramsperger, U.; Pohl, T.; Wood, V.; Rajandream, M.A.; Barrell, B.G.  
 submitted to the EMBL Data Library, October 1999  
 A:Reference number: Z21952  
 A:Accession: T40856  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-572 <RAM>  
 A:Cross-references: EMBL:AL121783; PIDN: CAB57847.1; GSPDB: GN00068; SPDB: SPCC11E10.05c  
 A:Experimental source: strain 972h-; cosmid c11E10  
 C:Genetics:  
 A:Gene: SPDB: SPCC11E10.05c  
 A:Map position: 3

Query Match 15.4%; Score 347; DB 2; Length 572;  
 Best Local Similarity 25.5%; Pred. No. 6.6e-21;  
 Matches 115; Conservative 67; Mismatches 157; Indels 112; Gaps 16;

Qy 49 YGIMFDAGSTGTRHVTTF-----VQKMPGQLPILGEVFD-----SVKPLSLAFVD 95  
 Db 5 YGIFIDAGSGSRLIYSWDYDTSLSKDKVKKLPLETGIGDGKWSLKVPQGISFAN 64  
 Qy 96 QPKQ-GAETVOGLEVAKDSIPRSHWKTTPVVLKATAGLRLL-PEHKAKALLPEVKIPE 153  
 Db 65 NPKHVKKHLKELLDFAHAIPKDVHETVPVLSATAGMLLGVDAONKILSHACRYIKK 124  
 Qy 154 KSPFLVP--KGSVIMDSGEGILAWTVNFLTQQLHGRQETVGTLDLGAASQTITFLP 211  
 Db 125 NYDFDIENCNSIRVIDGKAGMYGLATNVLNLTLEEKDTSTVGLDMGASVQIAF-- 182  
 Qy 212 QFEKTLQTRPGYLTSEMENST-----YKLYTHSYLGLFGKLAARLATLGL- 258  
 Db 183 -----ELPPSQLNKYKDSISTVHIGLONQQQLYFLPFTVTLWLGFGANEAYRYLLGLLI 235  
 Qy 259 ETE-GTDGHTFRSACLPRWLEAEWIFGQVKYQYGG-----NOGEVGPPEPCY 304  
 Db 236 ESENGKVNTLSDPCSIR--GRYDIDGIEFAGTGLDKQCLKLTYNLLNKDKPCSMWPCN 293  
 Qy 305 AEVLVRVVGKLUHQPEEVQSGSYAFSYYYDRAVDTMDIYKGGILKVEDFERKAREVC- 363  
 Db 294 FDGISI-----PPVDPAFTEFVGVSEFWYTTNDV----FDMGSGYHFPNFKYKVDYCYG 343  
 Qy 364 -----DNLENFTSGSPFLCMDSYITALLKDGFGFADSTV----- 398  
 Db 344 TEWETMLSRLYNKELTSTDENKLEK-----LCFKAASWALNVLHEGFDVPPKSNSSND 396

Qy 399 -----LQTKKVNNIETGWALG 415  
 Db 397 AKDGLSVIPAYHSPFTSLEKIERTEVSWTLG 427

RESULT 13  
 E86276  
 hypothetical protein F14L17.1 - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Dec-2001  
 C:Accession: E86276  
 R:Rheologas, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
 Chin, N.F.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.  
 ansen, N.F.; Hughes, B.; Huizar, L.  
 Nature 408, 816-820, 2000  
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.;  
 C.A.; Li, J.H.; Li, Y.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Matti, R.; Marziali,  
 Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
 ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
 A:Reference number: A86141; MUID: 21016719; PMID: 11130712  
 A:Accession: E86276  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-405 <STO>  
 A:Cross-references: GB:AE005172; NID: g7262666; PIDN: AAF43924.1; GSPDB: GN00141  
 C:Genetics:  
 A:Map position: 1

Query Match 14.6%; Score 329.5; DB 2; Length 405;  
 Best Local Similarity 26.3%; Pred. No. 1.2e-19;  
 Matches 110; Conservative 63; Mismatches 158; Indels 87; Gaps 15;

Qy 48 LYGIMFDAGSTGTRHVTTFVQKMPGQLPILGEVFDVSKVPGLSAFVDQPKQAETVQGL 107  
 Db 3 VFGYWFESGKP-----VFDFGEHYASL-----KLSPGLSSYADNPEGASVSVTKL 48  
 Qy 108 LEVAKDSIPRSHWKTTPVVLKATAGLRLLPEHKAKALLPEVKIPEFRKSPFLVPKGSVIM 167  
 Db 49 VEFAGRIKPKGKUKSDIRLMATAGWRLLDVPVQELDVTTRVLRSSGFKQDEWATVI 108  
 Qy 168 DGSDEGLAWTVNFLTQQLHGRQETVGTLDLGAASQTITFLPQFTELETPRGYLT 227  
 Db 109 SGTDGIIYAWVANHALGSLGGDPLKTTGIVELGGASAQVTFVPS-----EHPVPEFSRT 163  
 Qy 228 FEMFNSTYKLYTHSYLGLFGKLAARLATLGALETE-----GTDGHTFRSACLPR- 275  
 Db 164 ISYGNVSYTYSHSFDFGQDAEADKLLSLSQNSVAASGTDGIVDPCTPKGYIYDTHSQ 223  
 Qy 276 -----WLEAEWIF-GGQVKYQYGGNQEVEGVEPCYAEVLVRVVR-GKLUHQPEEVQSGSYA 328  
 Db 224 KDSGGLSESKPKASLQVQAAGD-----FTKRSATLMLQEGKEN----- 265  
 Qy 329 FSYYYDRAVDTMDIYKGGILKVEDF-----ERKAREVCNKL 366  
 Db 266 -CAYKHCISGSTTPNIQSFLENFFHTSKFGLGEKLEWLESEMILAKRFGCEWSKL 324  
 Qy 367 -ENF-TSGSPFL---CMDSYITALLKDGFGFA-DSTVLQTLTKV--NNIETGWALGA 416  
 Db 325 KEKYPTTKDKYLHRYCFSSAYIISMLHSLGVALDDEIRIKYASKAGKENIPLDVALGA 382

RESULT 14  
 T04439  
 hypothetical protein T18B16.150 - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 30-Apr-1999 #sequence\_revision 30-Apr-1999 #text\_change 30-Apr-1999  
 C:Accession: T04439  
 R:Bevan, M.; Benes, V.; Rechmann, S.; Borkova, D.; Anson, W.; Bancroft, I.; Mewes, H.W.  
 submitted to the Protein Sequence Database, April 1998  
 A:Reference number: Z15359

A;Accession: T04439

A;Molecule type: DNA

A;Residues: 1-1052 &lt;BEV&gt;

A;Cross-references: EMBL:AL021687

A;Experimental source: cultivar Columbia; BAC clone T18B16

C;Genetics:

A;Map position: 4

A;Introns: 79/3; 118/3; 217/1; 295/3; 396/3; 531/2; 815/3

A;Note: T18B16.150

Query Match 12.2%; Score 274; DB 2; Length 1052;

Best Local Similarity 24.9%; Pred. No. 2.1e-14;

Matches 107; Conservative 70; Mismatches 171; Indels 82; Gaps 19;

QY 51 IMPDAGSTGRIHVY--TFQKMPGOLPIL-----RGEVFD--SVKPGLSAFV 94

DB 525 LVIVISITGTRAVYQASINYNKQSSLPVWKSITGSRKSGRAYDRMETEPGDKLV 594

QY 95 DOPKOAGET--VOGLEVAKDSIPRSHWKTTPVVLKATAGLRLPEHAKALLFEVKEIFR 153

DB 585 NN-RTGLKTAIKPLQWAEKQIPKNAHRTTSLFVYATAGVRRURPADSSWILGNVWSILA 643

QY 154 KSPFLVPKGSVSMDSDEGILAWTVNFLTQGLHG-HRQETVGTLDLGGASTQITFLPQ 212

DB 644 KSPFTCRREWKIISGTEEAYFGWTALNYQTSMLGALPKKATFGALDLGGSSLOVTFENE 703

QY 213 FEXTLQTPRGYLTSPFMENSTYKLYTHSYLGFGLKAA-----RLATL----- 255

DB 704 -EKTHNETN-----LNLRIGSVNHLGAYSLAGLNDADFRRSVVHLKLPVWKSDLIE 758

QY 256 GALETEGTGHTFRSACLPRWLEAEWIFGVKYQYGNQGEVGFPE-----PCYAEVLR 309

DB 759 GKLE-----MKHPLNSGVNGQYICQSCASSVQGGKGGKSVIKLVGAENWGECSA 810

QY 310 VVR-GKLUHQPEEVQR--GSPYASYDDRAVDTMDIDYEKGI-----LKVEDFERKAREVC 363

DB 811 LAKNAPCALPDGPRPHQGFYAVSGFF-----VYREFNLSAEASLDVLEKGREFC 862

QY 364 DN---LENFTSGSPFL---CMDLSYITALLKDGFGFADSTV-----LQLTKVNNIET 410

DB 863 DKAWQVARTSVSPQPIEQCFRAPHYVLSLREGLYITDKQIIIGSGSITWTGLVALLES 922

QY 411 GWALGATFHL 420

DB 923 GKALSSTLGL 932

## RESULT 15

C86276

7A19.33 protein - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 18-Nov-2002

C;Accession: C86276

R;Theologos, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,

Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;

ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A;Reference number: A86141; MUID:21016719; PMID:11130712

A;Accession: C86276

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-508 &lt;STO&gt;

A;Cross-references: GB:AE005172; NID:95080800; PIDN:AAD39310.1; GSPDB:GN00141

C;Genetics:

A;Map position: 1

C;Superfamily: nucleoside triphosphatase chromatin-associated

Query Match 11.0%; Score 248; DB 2; Length 508;  
Best Local Similarity 23.3%; Pred. No. 1.1e-12;  
Matches 110; Conservative 77; Mismatches 159; Indels 126; Gaps 21;

QY 12 LVVSCVCSAV-----SHRNOQTWFEIGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

DB 35 VIVACVTIAGLLFTGYILRSGRNR-----VSLHYSVIIDGSSGT 77

QY 61 RIHVTVFVQKMPGOLPILE--GEVFD--PGLSAPVDQPKQAGTVOGLEVAKDSIP 116

DB 78 RVHVEGY--RIESGKPVDFGEEYASLKLSPGLSAYADNPEGVSSTVTELVFAKRVH 135

QY 117 RSHWKTTPVVLKATAGLRLPEHAKALLFEVKEIFRSPFLVPKGSVSMDSDE--GI 174

DB 136 KGKLLKSDIRLMATAGMRLLLELPVQEQILDVTRVLRSSGF-----DPRDEWASV 185

QY 175 LAWTVNFLTQGLHGHROETVGTLDL-----GASTOITFLPQPEKTLQTPRGYLTGF 228

DB 186 ISEILENF-----QDMKVYMLGLLIMRSVRLEVTFFV-----STELVSEFSRTL 231

QY 229 EMFNSTYKLYTHSYLGFGLKAAARLATL-----ETEGTDGHTFRSACLPRWLEAEWI--- 282

DB 232 AYGNSVYNLYSHFLDFGODAAQEKLSLSLYNSAANSTGEGIVPDPICIPKGVILETNLQK 291

QY 283 -----FGVKYQYGNQGEVGFPEPCYAEVLRVVR-----GKLUHQPEEVQRGS 325

DB 292 DLPGLADKGTATLQAAGNFSECRSAFAMLOBEKGKCTYKRCSIGSIPTN--LOGS 349

QY 326 FYAFSVY-----DRAVDTMI-----DYKGGILKVEDPERKAREVCNDLE 367

DB 350 FLATENFFHTSKFGLGEKEWJSEMILAGKRCFGGEWSK---LKVKYPTFKD-----ENLL 402

QY 368 NFTSGSPFLCMDLSYITALLKDGFGFA-DSTVLQTLTKV--NNIETGAWALGA 416

DB 403 RV-----CFSSAYIIISMLHDSLGVALLDDERIKYASKAGESDIPLDWALGA 447

Search completed: January 29, 2004, 10:06:16

Job time : 19.5 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2004, CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 09:35:09 ; Search time 9.5 Seconds  
(without alignments)  
2118.677 Million cell updates/sec

Title: US-10-091-085-3

Perfect score: 2250

Sequence: 1 MATSWGVFVFLVVCVCSA.....ETGWALGATFHLLQSLGISH 428

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2250	100.0	428	1 ENP5_HUMAN	O75356 homo sapien
2	1994.5	88.6	427	1 ENP5_MOUSE	Q9wuz9 mus musculus
3	1933.5	85.9	469	1 ENP5_MESAU	O9qvc8 mesocricetu
4	999	44.4	484	1 ENP6_HUMAN	O75354 homo sapien
5	986	43.8	455	1 ENP6_RAT	O9ex31 rattus norv
6	489	21.7	455	1 NTPA_PEA	P52914 pisum sativ
7	483.5	21.5	518	1 GDAL_YEAST	P32621 saccharomyc
8	446.5	19.8	454	1 APY_SOLFU	P80595 solanum tub
9	387.5	17.2	485	1 YY4E_CABEL	Q18411 caenorhabdi
10	385	17.1	552	1 YBU4_CABEL	Q21815 caenorhabdi
11	383	17.0	493	1 ENP1_CHICK	O93295 gallus gall
12	372	16.5	630	1 YND1_YEAST	P40009 saccharomyc
13	370	16.4	510	1 ENP1_MOUSE	P55772 mus musculus
14	365	16.2	513	1 ENP1_BOVIN	O18956 bos taurus
15	360.5	16.0	613	1 ENP4_MOUSE	O9dbt4 mus musculus
16	357	15.9	616	1 ENP4_HUMAN	O9y227 homo sapien
17	352	15.6	510	1 ENP1_PIG	O9myv4 sus scrofa
18	351	15.6	510	1 ENP1_HUMAN	P49561 homo sapien
19	348.5	15.5	511	1 ENP1_RAT	P97687 rattus norv
20	347.5	15.4	529	1 ENP3_HUMAN	O75355 homo sapien
21	339.5	15.1	495	1 ENP2_HUMAN	O9y513 homo sapien
22	309.5	13.8	495	1 ENP2_MOUSE	O55026 mus musculus
23	302.5	13.4	495	1 ENP2_RAT	O35795 rattus norv
24	301.5	13.4	494	1 ENP2_CHICK	P79784 gallus gall
25	140	6.2	628	1 NTP1_TOXGO	Q27893 toxoplasma
26	138	6.1	628	1 NTP2_TOXGO	Q27895 toxoplasma
27	131	5.8	592	1 NTP3_TOXGO	P52913 toxoplasma
28	111.5	5.0	634	1 SELB_MOOTH	O46455 moorella th
29	106.5	4.7	553	1 VGI_SpV4	P11333 pipitoplasma
30	104.5	4.6	4427	1 PKSL_BACSU	O05470 bacillus su
31	101.5	4.5	1151	1 Y243_TREPA	O83273 treponema p
32	99.5	4.4	535	1 DEXB_STRPN	O54796 streptococ
33	99.5	4.4	548	1 THER_BACST	P06874 bacillus st

34	97	4.3	1220	1	PTC1_BRARE	Q98864 brachydanio
35	96	4.3	711	1	DPE3_YEAST	Q08225 saccharomyc
36	95.5	4.2	2076	1	FAS1_YARLI	P34229 y fatty aci
37	93	4.1	969	1	PAC4_HUMAN	P29122 homo sapien
38	93	4.1	1696	1	ITN2_HUMAN	Q9nzm3 homo sapien
39	92	4.1	290	1	HEN6_CAUCR	Q9ast8 caulobacter
40	91	4.0	417	1	GLYA_BORBU	O51947 borrelia bu
41	91	4.0	477	1	URTI_DESRO	P98119 desmodus ro
42	90	4.0	434	1	YFEW_ECO57	Q8xbj0 escherichia
43	90	4.0	434	1	YFEW_ECOLI	P77619 escherichia
44	89.5	4.0	809	1	RS95_YEAST	P39940 saccharomyc
45	89.5	4.0	2334	1	WAPA_BACSU	Q07833 bacillus su

## ALIGNMENTS

RESULT 1	ENP5_HUMAN	STANDARD;	PRT;	428 AA.
ID	ENP5_HUMAN			
AC	O75356;			
DT	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DT	16-OCT-2001 (Rel. 40, Last annotation update)			
DE	Ectonucleoside triphosphate diphosphohydrolase 5 precursor			
DE	(EC 3.6.1.6) (NTPDase5) (Nucleoside diphosphatase) (CD39 antigen-like			
DE	4) (ER-UDPase);			
GN	ENTPD5 OR CD39L4.			
OS	Homo sapiens (Human)			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Leukemia;			
RX	MEDLINE=98341119; PubMed=9676430;			
RA	Chadwick B.P., Frischauf A.-M.;			
RT	"The CD39-like gene family: identification of three new human members			
RT	(CD39L2, CD39L3, and CD39L4), their murine homologues, and a member of			
RT	the gene family from Drosophila melanogaster.";			
RL	Genomics 50:357-367(1998)			
CC	-!- FUNCTION: LIKELY TO PROMOTE REGLYCOSYLATION REACTIONS INVOLVED IN			
CC	GLYCOPROTEINS FOLDING AND QUALITY CONTROL IN THE ENDOPLASMIC			
CC	RETICULUM. HYDROLYZES UDP, GDP AND IDP BUT NOT ANY OTHER			
CC	NUCLEOSIDE DI-, MONO- OR TRIPHOSPHATES, NOR THIAMINE PYROPHOSPHATE			
CC	(BY SIMILARITY).			
CC	-!- CATALYTIC ACTIVITY: A nucleoside diphosphate + H(2)O = a			
CC	nucleotide + phosphate.			
CC	-!- COFACTOR: REQUIRES CALCIUM AND MAGNESIUM.			
CC	-!- SUBCELLULAR LOCATION: Endoplasmic reticulum lumen (By similarity).			
CC	-!- TISSUE SPECIFICITY: EXPRESSED IN ADULT LIVER, KIDNEY, PROSTATE,			
CC	TESTIS AND COLON. MUCH WEAKER EXPRESSION IN OTHER TISSUES.			
CC	-!- SIMILARITY: BELONGS TO THE GDAL / CD39 NTPASE FAMILY.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
CC	EMBL; AF039918; AAC39885.1; ..			
CC	Genew; HGNC:3367; ENTPD5.			
CC	MIM; 603162; ..			
CC	GO; GO:0004002; F:adenosinetriphosphatase activity; TAS.			
CC	InterPro; IPR000407; GDAL_CD39_NTPase.			
CC	Pfam; PF01150; GDAL_CD39.1.			
CC	PROSITE; PS01238; GDAL_CD39_NTPASE; FALSE NEG.			
CC	Hydrolase; Transmembrane; Glycoprotein; Calcium; Magnesium;			
CC	Endoplasmic reticulum; Signal.			
CC	SIGNAL 1 20 POTENTIAL.			
CC	CHAIN 21 428 ECTONUCLEOSIDE TRIPHOSPHATE			

FT CARBOHYD 42 42 DIPHOSEPHOHYDROLASE 5.  
 FT CARBOHYD 232 232 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 368 368 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 428 AA; 47517 MW; 830437A155D4DDD CRC64;  
 Query Match 100.0%; Score 2250; DB 1; Length 428;  
 Best Local Similarity 100.0%; Pred. No. 3.3e-173;  
 Matches 428; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 MATSGVTFVFMVSVCSAVSHRNQQTWTFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60  
 Db 1 MATSGVTFVFMVSVCSAVSHRNQQTWTFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60  
 Qy 61 RIHYVTFVQKMPGOLPILGEGVDFSVKPLGSAFVDQPKQAGTVOGLEVAKDSIPRSHW 120  
 Db 61 RIHYVTFVQKMPGOLPILGEGVDFSVKPLGSAFVDQPKQAGTVOGLEVAKDSIPRSHW 120  
 Qy 121 KKTVPVLKATAGLRLLPEHKAALLFEVKEIFRKSPLVPKGSVIMDGSDEGILAWTV 180  
 Db 121 KKTVPVLKATAGLRLLPEHKAALLFEVKEIFRKSPLVPKGSVIMDGSDEGILAWTV 180  
 Qy 181 NFLTQGLHGHQRTVGTLDLGGASTQITFLPQPKETLEQTPRGYLTSPFEMFNSTYKLYTH 240  
 Db 181 NFLTQGLHGHQRTVGTLDLGGASTQITFLPQPKETLEQTPRGYLTSPFEMFNSTYKLYTH 240  
 Qy 241 SYLFGGLKAARLATLGALETSGTDGHTFRSACLPRWLEAEWIFGVKYQYGNQGEVGF 300  
 Db 241 SYLFGGLKAARLATLGALETSGTDGHTFRSACLPRWLEAEWIFGVKYQYGNQGEVGF 300  
 Qy 301 EPCYAEVLVRVVRKGLHQPVEEVQVGSFYAFSYYVDRAVDTMDIDYKGGILKVEDFERKAR 360  
 Db 301 EPCYAEVLVRVVRKGLHQPVEEVQVGSFYAFSYYVDRAVDTMDIDYKGGILKVEDFERKAR 360  
 Qy 361 EVCNLENFTSGSPFLCMLDLYITALLKDGFGFADSTVLQTKKNNIETGALGATFHL 420  
 Db 361 EVCNLENFTSGSPFLCMLDLYITALLKDGFGFADSTVLQTKKNNIETGALGATFHL 420  
 Qy 421 LOSLGISH 428  
 Db 421 LOSLGISH 428  
 RESULT 2  
 ENP5\_MOUSE STANDARD; PRT; 427 AA.  
 AC Q9WU29; O70214;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Ectonucleoside triphosphate diphosphohydrolase 5 precursor  
 DE (EC 3.6.1.6) (NTPDase5) (Nucleoside diphosphatase) (CD39 antigen-like  
 DE 4) (ER-UDPase).  
 GN ENTPD5 OR CD39L4.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_taxid=10090;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RC TISSUE=Testis;  
 RX MEDLINE=98119025; PubMed=9457681;  
 RA Chadwick B.P., Williamson J., Sheer D., Frischaut A.-M.;  
 RT "cDNA cloning and chromosomal mapping of a mouse gene with homology to  
 NTPases.";  
 RL Mamm. Genome 9:162-164 (1998) -  
 RN [2]  
 RP SEQUENCE FROM N.A., SUBCELLULAR LOCATION, AND CHARACTERIZATION.  
 RC TISSUE=Liver;  
 RX MEDLINE=99298181; PubMed=10369669;  
 RA Trombetta E.S., Helenius A.;  
 RT "Glycoprotein reglucosylation and nucleotide sugar utilization in the  
 secretory pathway: identification of a nucleoside diphosphatase in the

endoplasmic reticulum.";  
 EMBO J. 18:3282-3292 (1999).  
 [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Kidney;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
 RA Saito T., Okazaki Y., Gotohori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,  
 RA Schirml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustinich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whittaker C., Wilming L.,  
 RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohtsuki S.,  
 RA Hayashizaki Y.;  
 RT "Functional annotation of a full-length mouse cDNA collection.";  
 RL Nature 409:685-690 (2001).  
 CC -1- FUNCTION: TO PROMOTE REGLYCOSYLATION REACTIONS INVOLVED IN  
 CC GLYCOPROTEIN FOLDING AND QUALITY CONTROL IN THE ENDOPLASMIC  
 CC RETICULUM. HYDROLYZES UDP, GDP AND IDP BUT NOT ANY OTHER  
 CC NUCLEOSIDE DI-, MONO- OR TRIPHOSPHATES, NOR THIAMINE  
 CC PYROPHOSPHATE.  
 CC -1- CATALYTIC ACTIVITY: A nucleoside diphosphate + H(2)O = a  
 CC nucleotide + phosphate.  
 CC -1- COFACTOR: REQUIRES CALCIUM AND MAGNESIUM.  
 CC -1- SUBCELLULAR LOCATION: Endoplasmic reticulum lumen.  
 CC -1- TISSUE SPECIFICITY: Ubiquitous.  
 CC -1- PFM: GLYCOSYLATED WITH HIGH MANNOSE N-LINKED GLYCANS.  
 CC -1- MISCELLANEOUS: OPTIMAL PH IS NEUTRAL.  
 CC -1- SIMILARITY: BELONGS TO THE GDAI / CD39 NTPASE FAMILY.  
 CC -----  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AF006482; AAC05181.1; -  
 CC EMBL; AJ238636; CAB45533.1; -  
 CC EMBL; AK002618; BAB22234.1; -  
 CC MCD; MG1.1321385; Entdgs.  
 CC InterPro; IPR000407; GDAI\_CD39\_NTPase.  
 CC Pfam; PF01150; GDAI\_CD39; 1.  
 CC PROSITE; PS01238; GDAI\_CD39\_NTPASE; FALSE NEG.  
 CC Hydrolase; Transmembrane; Glycoprotein; Calcium; Magnesium;  
 CC Endoplasmic reticulum; Signal.  
 FT SIGNAL 1 18  
 FT CHAIN 19 427  
 FT ECTONUCLEOSIDE TRIPHOSPHATE  
 FT DIPHOSPHOHYDROLASE 5.  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 41 41  
 FT CARBOHYD 231 231  
 FT CONFLICT 390 390  
 FT CONFLICT 394 427  
 FT SIGNAL 1 18  
 FT CHAIN 19 427  
 FT ECTONUCLEOSIDE TRIPHOSPHATE  
 FT DIPHOSPHOHYDROLASE 5.  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 41 41  
 FT CARBOHYD 231 231  
 FT CONFLICT 390 390  
 FT CONFLICT 394 427  
 FT SIGNAL 1 18  
 FT CHAIN 19 427  
 FT ECTONUCLEOSIDE TRIPHOSPHATE  
 FT DIPHOSPHOHYDROLASE 5.  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 41 41  
 FT CARBOHYD 231 231  
 FT CONFLICT 390 390  
 FT CONFLICT 394 427  
 FT SIGNAL 1 18  
 FT CHAIN 19 427  
 FT ECTONUCLEOSIDE TRIPHOSPHATE  
 FT DIPHOSPHOHYDROLASE 5.  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 41 41  
 FT CARBOHYD 231 231  
 FT CONFLICT 390 390  
 FT CONFLICT 394 427  
 FT SIGNAL 1 18  
 FT CHAIN 19 427  
 FT ECTONUCLEOSIDE TRIPHOSPHATE  
 FT DIPHOSPHOHYDROLASE 5.  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 41 41  
 FT CARBOHYD 231 231  
 FT CONFLICT 390 390  
 FT CONFLICT 394 427  
 FT SIGNAL 1 18  
 FT CHAIN 19 427  
 FT ECTONUCLEOSIDE TRIPHOSPHATE  
 FT DIPHOSPHOHYDROLASE 5.  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 41 41  
 FT CARBOHYD 231 231  
 FT CONFLICT 390 390  
 FT CONFLICT 394 427  
 FT SIGNAL 1 18  
 FT CHAIN 19 427  
 FT ECTONUCLEOSIDE TRIPHOSPHATE  
 FT DIPHOSPHOHYDROLASE 5.  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 41 41  
 FT CARBOHYD 231 231  
 FT CONFLICT 390 390  
 FT CONFLICT 394 427  
 FT SIGNAL 1 18  
 FT CHAIN 19 427  
 FT ECTONUCLEOSIDE TRIPHOSPHATE  
 FT DIPHOSPHOHYDROLASE 5.  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 41 41  
 FT CARBOHYD 231 231  
 FT CONFLICT 390 390  
 FT CONFLICT 394 427  
 FT SIGNAL 1 18  
 FT CHAIN 19 427  
 FT ECTONUCLEOSIDE TRIPHOSPHATE  
 FT DIPHOSPHOHYDROLASE 5.  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 41 41  
 FT CARBOHYD 231 231  
 FT CONFLICT 390 390  
 FT CONFLICT 394 427  
 FT SIGNAL 1 18  
 FT CHAIN 19 427  
 FT ECTONUCLEOSIDE TRIPHOSPHATE  
 FT DIPHOSPHOHYDROLASE 5.  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 41 41  
 FT CARBOHYD 231 231  
 FT CONFLICT 390 390  
 FT CONFLICT 394 427  
 FT SIGNAL 1 18  
 FT CHAIN 19 427  
 FT ECTONUCLEOSIDE TRIPHOSPHATE  
 FT DIPHOSPHOHYDROLASE 5.  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 41 41  
 FT CARBOHYD 231 231  
 FT CONFLICT 390 390  
 FT CONFLICT 394 427  
 FT SIGNAL 1 18  
 FT CHAIN 19 427  
 FT ECTONUCLEOSIDE TRIPHOSPHATE  
 FT DIPHOSPHOHYDROLASE 5.  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 41 41  
 FT CARBOHYD 231 231  
 FT CONFLICT 390 390  
 FT CONFLICT 394 427  
 FT SIGNAL 1 18  
 FT CHAIN 19 427  
 FT ECTONUCLEOSIDE TRIPHOSPHATE  
 FT DIPHOSPHOHYDROLASE 5.  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 41 41  
 FT CARBOHYD 231 231  
 FT CONFLICT 390 390  
 FT CONFLICT 394 427  
 FT SIGNAL 1 18  
 FT CHAIN 19 427  
 FT ECTONUCLEOSIDE TRIPHOSPHATE  
 FT DIPHOSPHOHYDROLASE 5.  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 41 41  
 FT CARBOHYD 231 231  
 FT CONFLICT 390 390  
 FT CONFLICT 394 427  
 FT SIGNAL 1 18  
 FT CHAIN 19 427  
 FT ECTONUCLEOSIDE TRIPHOSPHATE  
 FT DIPHOSPHOHYDROLASE 5.  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 41 41  
 FT CARBOHYD 231 231  
 FT CONFLICT 390 390  
 FT CONFLICT 394 427  
 FT SIGNAL 1 18  
 FT CHAIN 19 427  
 FT ECTONUCLEOSIDE TRIPHOSPHATE  
 FT DIPHOSPHOHYDROLASE 5.  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 41 41  
 FT CARBOHYD 231 231  
 FT CONFLICT 390 390  
 FT CONFLICT 394 427  
 FT SIGNAL 1 18  
 FT CHAIN 19 427  
 FT ECTONUCLEOSIDE TRIPHOSPHATE  
 FT DIPHOSPHOHYDROLASE 5.  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 41 41  
 FT CARBOHYD 231 231  
 FT CONFLICT 390 390  
 FT CONFLICT 394 427  
 FT SIGNAL 1 18  
 FT CHAIN 19 427  
 FT ECTONUCLEOSIDE TRIPHOSPHATE  
 FT DIPHOSPHOHYDROLASE 5.  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 41 41  
 FT CARBOHYD 231 231  
 FT CONFLICT 390 390  
 FT CONFLICT 394 427  
 FT SIGNAL 1 18  
 FT CHAIN 19 427  
 FT ECTONUCLEOSIDE TRIPHOSPHATE  
 FT DIPHOSPHOHYDROLASE 5.  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 41 41  
 FT CARBOHYD 231 231  
 FT CONFLICT 390 390  
 FT CONFLICT 394 427  
 FT SIGNAL 1 18  
 FT CHAIN 19 427  
 FT ECTONUCLEOSIDE TRIPHOSPHATE  
 FT DIPHOSPHOHYDROLASE 5.  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 41 41  
 FT CARBOHYD 231 231  
 FT CONFLICT 390 390  
 FT CONFLICT 394 427  
 FT SIGNAL 1 18  
 FT CHAIN 19 427  
 FT ECTONUCLEOSIDE TRIPHOSPHATE  
 FT DIPHOSPHOHYDROLASE 5.  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 41 41  
 FT CARBOHYD 231 231  
 FT CONFLICT 390 390  
 FT CONFLICT 394 427  
 FT SIGNAL 1 18  
 FT CHAIN 19 427  
 FT ECTONUCLEOSIDE TRIPHOSPHATE  
 FT DIPHOSPHOHYDROLASE 5.  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 41 41  
 FT CARBOHYD 231 231  
 FT CONFLICT 390 390  
 FT CONFLICT 394 427  
 FT SIGNAL 1 18  
 FT CHAIN 19 427  
 FT ECTONUCLEOSIDE TRIPHOSPHATE  
 FT DIPHOSPHOHYDROLASE 5.  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 41 41  
 FT CARBOHYD 231 231  
 FT CONFLICT 390 390  
 FT CONFLICT 394 427  
 FT SIGNAL 1 18  
 FT CHAIN 19 427  
 FT ECTONUCLEOSIDE TRIPHOSPHATE  
 FT DIPHOSPHOHYDROLASE 5.  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 41 41  
 FT CARBOHYD 231 231  
 FT CONFLICT 390 390  
 FT CONFLICT 394 427  
 FT SIGNAL 1 18  
 FT CHAIN 19 427  
 FT ECTONUCLEOSIDE TRIPHOSPHATE  
 FT DIPHOSPHOHYDROLASE 5.  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 41 41  
 FT CARBOHYD 231 231  
 FT CONFLICT 390 390  
 FT CONFLICT 394 427  
 FT SIGNAL 1 18  
 FT CHAIN 19 427  
 FT ECTONUCLEOSIDE TRIPHOSPHATE  
 FT DIPHOSPHOHYDROLASE 5.  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 41 41  
 FT CARBOHYD 231 231  
 FT CONFLICT 390 390  
 FT CONFLICT 394 427  
 FT SIGNAL 1 18  
 FT CHAIN 19 427  
 FT ECTONUCLEOSIDE TRIPHOSPHATE  
 FT DIPHOSPHOHYDROLASE 5.  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 41 41  
 FT CARBOHYD 231 231  
 FT CONFLICT 390 390  
 FT CONFLICT 394 427  
 FT SIGNAL 1 18  
 FT CHAIN 19 427  
 FT ECTONUCLEOSIDE TRIPHOSPHATE  
 FT DIPHOSPHOHYDROLASE 5.  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 41 41  
 FT CARBOHYD 231 231  
 FT CONFLICT 390 390  
 FT CONFLICT 394 427  
 FT SIGNAL 1 18  
 FT CHAIN 19 427  
 FT ECTONUCLEOSIDE TRIPHOSPHATE  
 FT DIPHOSPHOHYDROLASE 5.  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 41 41  
 FT CARBOHYD 231 231  
 FT CONFLICT 390 390  
 FT CONFLICT 394 427  
 FT SIGNAL 1 18  
 FT CHAIN 19 427  
 FT ECTONUCLEOSIDE TRIPHOSPHATE  
 FT DIPHOSPHOHYDROLASE 5.  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 41 41  
 FT CARBOHYD 231 231  
 FT CONFLICT 390 390  
 FT CONFLICT 394 427  
 FT SIGNAL 1 18  
 FT CHAIN 19 427  
 FT ECTONUCLEOSIDE TRIPHOSPHATE  
 FT DIPHOSPHOHYDROLASE 5.  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 41 41  
 FT CARBOHYD 231 231  
 FT CONFLICT 390 390  
 FT CONFLICT 394 427  
 FT SIGNAL 1 18  
 FT CHAIN 19 427  
 FT ECTONUCLEOSIDE TRIPHOSPHATE  
 FT DIPHOSPHOHYDROLASE 5.  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 41 41  
 FT CARBOHYD 231 231  
 FT CONFLICT 390 390  
 FT CONFLICT 394 427  
 FT SIGNAL 1 18  
 FT CHAIN 19 427  
 FT ECTONUCLEOSIDE TRIPHOSPHATE  
 FT DIPHOSPHOHYDROLASE 5.  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 41 41  
 FT CARBOHYD 231 231  
 FT CONFLICT 390 390  
 FT CONFLICT 394 427  
 FT SIGNAL 1 18  
 FT CHAIN 19 427  
 FT ECTONUCLEOSIDE TRIPHOSPHATE  
 FT DIPHOSPHOHYDROLASE 5.  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 41 41  
 FT CARBOHYD 231 231  
 FT CONFLICT 390 390  
 FT CONFLICT 394 427  
 FT SIGNAL 1 18  
 FT CHAIN 19 427  
 FT ECTONUCLEOSIDE TRIPHOSPHATE  
 FT DIPHOSPHOHYDROLASE 5.  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 41 41  
 FT CARBOHYD 231 231  
 FT CONFLICT 390 390  
 FT CONFLICT 394 427  
 FT SIGNAL 1 18  
 FT CHAIN 19 427  
 FT ECTONUCLEOSIDE TRIPHOSPHATE  
 FT DIPHOSPHOHYDROLASE 5.  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 41 41  
 FT CARBOHYD 231 231  
 FT CONFLICT 390 390  
 FT CONFLICT 394 427  
 FT SIGNAL 1 18  
 FT CHAIN 19 427  
 FT ECTONUCLEOSIDE TRIPHOSPHATE  
 FT DIPHOSPHOHYDROLASE 5.  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 41 41  
 FT CARBOHYD 231 231  
 FT CONFLICT 390 390  
 FT CONFLICT 394 427  
 FT SIGNAL 1 18  
 FT CHAIN 19 427  
 FT ECTONUCLEOSIDE TRIPHOSPHATE  
 FT DIPHOSPHOHYDROLASE 5.  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 41 41  
 FT CARBOHYD 231 231  
 FT CONFLICT 390 390  
 FT CONFLICT 394 427  
 FT SIGNAL 1 18  
 FT CHAIN 19 427  
 FT ECTONUCLEOSIDE TRIPHOSPHATE  
 FT DIPHOSPHOHYDROLASE 5.  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 41 41  
 FT CARBOHYD 231 231  
 FT CONFLICT 390 390  
 FT CONFLICT 394 427  
 FT SIGNAL 1 18  
 FT CHAIN 19 427  
 FT ECTONUCLEOSIDE TRIPHOSPHATE  
 FT DIPHOSPHOHYDROLASE 5.  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 41 41  
 FT CARBOHYD 231 231  
 FT CONFLICT 390 390  
 FT CONFLICT 394 427  
 FT SIGNAL 1 18  
 FT CHAIN 19 427  
 FT ECTONUCLEOSIDE TRIPHOSPHATE  
 FT DIPHOSPHOHYDROLASE 5.  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 41 41  
 FT CARBOHYD 231 231  
 FT CONFLICT 390 390  
 FT CONFLICT 394 427  
 FT SIGNAL 1 18  
 FT CHAIN 19 427  
 FT ECTONUCLEOSIDE TRIPHOSPHATE  
 FT DIPHOSPHOHYDROLASE 5.  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 41 41  
 FT CARBOHYD 231 231  
 FT CONFLICT 390 390  
 FT CONFLICT 394 427  
 FT SIGNAL 1 18  
 FT CHAIN 19 427  
 FT ECTONUCLEOSIDE TRIPHOSPHATE  
 FT DIPHOSPHOHYDROLASE 5.  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 41 41  
 FT CARBOHYD 231 231  
 FT CONFLICT 390 390  
 FT CONFLICT 394 427  
 FT SIGNAL 1 18  
 FT CHAIN 19 427  
 FT ECTONUCLEOSIDE TRIPHOSPHATE  
 FT DIPHOSPHOHYDROLASE 5.  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 41 41  
 FT CARBOHYD 231 231  
 FT CONFLICT 390 390  
 FT CONFLICT 394 427  
 FT SIGNAL 1 18  
 FT CHAIN 19 427  
 FT ECTONUCLEOSIDE TRIPHOSPHATE  
 FT DIPHOSPHOHYDROLASE 5.  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 41 41  
 FT CARBOHYD 231 231  
 FT CONFLICT 390 390  
 FT CONFLICT 394 427  
 FT SIGNAL 1 18  
 FT CHAIN 19 427  
 FT ECTONUCLEOSIDE TRIPHOSPHATE  
 FT DIPHOSPHOHYDROLASE 5.  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 41 41  
 FT CARBOHYD 231 231  
 FT CONFLICT 390 390  
 FT CONFLICT 394 427  
 FT SIGNAL 1 18  
 FT CHAIN 19 427  
 FT ECTONUCLEOSIDE TRIPHOSPHATE  
 FT DIPHOSPHOHYDROLASE 5.  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 41 41  
 FT CARBOHYD 231 231  
 FT CONFLICT 390 390  
 FT CONFLICT 394 427  
 FT SIGNAL 1 18  
 FT CHAIN 19 427  
 FT ECTONUCLEOSIDE TRIPHOSPHATE  
 FT DIPHOSPHOHYDROLASE 5.  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 41 41  
 FT CARBOHYD 231 231  
 FT CONFLICT 390 390  
 FT CONFLICT 394 427  
 FT SIGNAL 1 18  
 FT CHAIN 19 427  
 FT ECTONUCLEOSIDE TRIPHOSPHATE  
 FT DIPHOSPHOHYDROLASE 5.  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 41 41  
 FT CARBOHYD 231 231  
 FT CONFLICT 390 390  
 FT CONFLICT 394 427  
 FT SIGNAL 1 18  
 FT CHAIN 19 427  
 FT ECTONUCLEOSIDE TRIPHOSPHATE  
 FT DIPHOSPHOHYDROLASE 5.  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 41 41  
 FT CARBOHYD 231 231  
 FT CONFLICT 390 390  
 FT CONFLICT 394 427  
 FT SIGNAL 1 18  
 FT CHAIN 19 427  
 FT ECTONUCLEOSIDE TRIPHOSPHATE  
 FT DIPHOSPHOHYDROLASE 5.  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 41 41  
 FT CARBOHYD 231 231  
 FT CONFLICT 390 390  
 FT CONFLICT 394 427  
 FT SIGNAL 1 18  
 FT CHAIN 19 427  
 FT ECTONUCLEOSIDE TRIPHOSPHATE  
 FT DIPHOSPHOHYDROLASE 5.  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 41 41  
 FT CARBOHYD 231 231  
 FT CONFLICT 390 390  
 FT CONFLICT 394 427  
 FT SIGNAL 1 18  
 FT CHAIN 19 427  
 FT ECTONUCLEOSIDE TRIPHOSPHATE  
 FT DIPHOSPHOHYDROLASE 5.  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 41 41  
 FT CARBOHYD 231 231  
 FT CONFLICT 390 390  
 FT CONFLICT 394 427  
 FT SIGNAL 1 18  
 FT CHAIN 19 427  
 FT ECTONUCLEOSIDE TRIPHOSPHATE  
 FT DIPHOSPHOHYDROLASE 5.  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 41 41  
 FT CARBOHYD 231 231  
 FT CONFLICT 390 390  
 FT CONFLICT 394 427  
 FT SIGNAL 1 18  
 FT CHAIN 19 427  
 FT ECTONUCLEOSIDE TRIPHOSPHATE  
 FT DIPHOSPHOHYDROLASE 5.  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 41 41  
 FT CARBOHYD 231 231  
 FT CONFLICT 390 390  
 FT CONFLICT 394 427  
 FT SIGNAL 1 18  
 FT CHAIN 19 427  
 FT ECTONUCLEOSIDE TRIPHOSPHATE  
 FT DIPHOSPHOHYDROLASE 5.  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 41 41  
 FT CARBOHYD 231 231  
 FT CONFLICT 390 390  
 FT CONFLICT 394 427  
 FT SIGNAL 1 18  
 FT CHAIN 19 427  
 FT ECTONUCLEOSIDE TRIPHOSPHATE  
 FT DIPHOSPHOHYDROLASE 5.  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 41 41  
 FT CARBOHYD 231 231  
 FT CONFLICT 390 390  
 FT CONFLICT 394 427  
 FT SIGNAL 1 18  
 FT CHAIN 19 427  
 FT ECTONUCLEOSIDE TRIPHOSPHATE  
 FT DIPHOSPHOHYDROLASE 5.  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 41 41  
 FT CARBOHYD 231 231  
 FT CONFLICT 390 390  
 FT CONFLICT 394 427  
 FT SIGNAL 1 18  
 FT CHAIN 19 427  
 FT ECTONUCLEOSIDE TRIPHOSPHATE  
 FT DIPHOSPHOHYDROLASE 5.  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 41 41  
 FT CARBOHYD 231 231  
 FT CONFLICT 390 390  
 FT CONFLICT 394 427  
 FT SIGNAL 1 18  
 FT CHAIN 19 427  
 FT ECTONUCLEOSIDE TRIPHOSPHATE  
 FT DIPHOSPHOHYDROLASE 5.  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 41 41  
 FT CARBOHYD 231 231  
 FT CONFLICT 390 390  
 FT CONFLICT 394 427  
 FT SIGNAL 1 18  
 FT CHAIN 19 427  
 FT ECTONUCLEOSIDE TRIPHOSPHATE  
 FT DIPHOSPHOHYDROLASE 5.  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 41 41  
 FT CARBOHYD 231 231  
 FT CONFLICT 390 390  
 FT CONFLICT 394 427  
 FT SIGNAL 1 18  
 FT CHAIN 19 427  
 FT ECTONUCLEOSIDE TRIPHOSPHATE  
 FT DIPHOSPHOHYDROLASE 5.  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 41 41  
 FT CARBOHYD 231 231  
 FT CONFLICT 390 390  
 FT CONFLICT 394 427  
 FT SIGNAL 1 18  
 FT CHAIN 19 427  
 FT ECTONUCLEOSIDE TRIPHOSPHATE  
 FT DIPHOSPHOHYDROLASE 5.  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 41 41  
 FT CARBOHYD 231 231  
 FT CONFLICT 390 390  
 FT CONFLICT 394 427  
 FT SIGNAL 1 18  
 FT CHAIN 19 427  
 FT ECTONUCLEOSIDE TRIPHOSPHATE  
 FT DIPHOSPHOHYDROLASE 5.  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 41 41  
 FT CARBOHYD 231 231  
 FT CONFLICT 390 390  
 FT CONFLICT 394 427  
 FT SIGNAL 1 18  
 FT CHAIN 19 427  
 FT ECTONUCLEOSIDE TRIPHOSPHATE  
 FT DIPHOSPHOHYDROLASE 5.  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).

```

Db 1 MATSGAV-FMLIIACVGTVFYRQQTWFEVFLSSMCPINVSAGTFYGMFDAGSTGT 59
Qy 61 RIHVYTFVOKMPQLPILGEVDSVKPGLSAPVDOPKOGAETVQGLLEVAKDSIPRSHW 120
Db 60 RIHVYTFVOKTAGQLPFLGEIEFDSVKPGLSAPVDOPKOGAETVQGLLEVAKDSIPRSHW 119
Qy 121 KKTVPVVKATAGLRLLEPHKAKALLFEVKEIFRKSPLFPKGSVIMDGSDEGILAWTV 180
Db 120 ERTFVVKATAGLRLLEPHKAKALLFEVKEIFRKSPLFPKGSVIMDGSDEGILAWTV 179
Qy 181 NFLTGQHGHRQBTGTDLGGASTQITFLPQPEKLEOTPRGYLTSFEMFNSTFKLYTH 240
Db 180 NFLTGQHGHRQBTGTDLGGASTQITFLPQPEKLEOTPRGYLTSFEMFNSTFKLYTH 239
Qy 241 SYLGFGKKAARLALATLGALETGTGHTFRSACLPRWLEAEWIFGGVKYQVGGNQEVEGF 300
Db 240 SYLGFGKKAARLALATLGALETGTGHTFRSACLPRWLEAEWIFGGVKYQVGGNQEVEGF 299
Qy 301 EPCYAEVLRVVRGKLHQPVEVQSGVAFYSYVYDRAVDTMDIDYKGGILKVEDFERKAR 360
Db 300 EPCYAEVLRVVRGKLHQPVEVQSGVAFYSYVYDRAVDTMDIDYKGGILKVEDFERKAR 359
Qy 361 EVCNLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQTKKNNIETGWALGATFHL 420
Db 360 EVCNLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQTKKNNIETGWALGATFHL 419
Qy 421 LQSLGTS 427
Db 420 LQSLGIT 426

RESULT 3
ENP6_HUMAN
ID ENP6_HUMAN STANDARD; PRT; 469 AA.
AC Q9QYC8; Q9QYC9;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ectonucleoside triphosphate diphosphohydrolase 5 precursor
DE (EC 3.6.1.6) (NTPases) (Nucleoside diphosphatase) (CD39 antigen-like
DE 4) (ER-UDPase) (Proto-oncogene cph).
GN Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
OX NCBI_TaxID=10036;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99142925; PubMed=9989819;
RA Velasco J.A., Avila M.A., Notario V.;
RT "The product of the cph oncogene is a truncated, nucleotide-binding
RT protein that enhances cellular survival to stress.";
RL Oncogene 18:689-701(1999).
CC -1- FUNCTION: LIKELY TO PROMOTE REGLYCOSYLATION REACTIONS INVOLVED IN
CC GLYCOPROTEIN FOLDING AND QUALITY CONTROL IN THE ENDOPLASMIC
CC RETICULUM. HYDROLYZES UDP, GDP AND IDP BUT NOT ANY OTHER
CC NUCLEOSIDE DI-, MONO- OR TRIPHOSPHATES, NOR THIAMINE PYROPHOSPHATE
CC (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: A nucleoside diphosphate + H(2)O = a
CC nucleotide + phosphate.
CC -1- COFACTOR: REQUIRES CALCIUM AND MAGNESIUM (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Endoplasmic reticulum lumen (By similarity).
CC -1- TISSUE SPECIFICITY: EXPRESSED IN FETAL CELLS AND MOST ADULT
CC TISSUES.
CC -1- DISEASE: CPH ONCOGENE HAS TRANSFORMING CAPACITY AND TUMORIGENIC
CC POTENTIAL.
CC -1- SIMILARITY: BELONGS TO THE GDA1 / CD39 NTPASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; AF084568; AAF22931.1; ALT_TERM.
CC EMBL; AF084569; AAF22932.1; --
CC InterPro; IPR000407; GDA1_CD39_NTPase.
CC Pfam; PF01150; GDA1_CD39; 1.
CC PROSITE; PS01238; GDA1_CD39_NTPASE; FALSE NEG.
KW Hydrolyase; Transmembrane; Glycoprotein; Calcium; Magnesium;
KW Endoplasmic reticulum; Signal; Proto-oncogene.
FT SIGNAL 1 18
FT CHAIN 19 469
FT CARBOHYD 42 42
FT CARBOHYD 232 232
FT SEQUENCE 469 AA; 52125 MW; 03D8A23B0C73474B CRC64;
SQ
Query Match 85.9%; Score 1933.5; DB 1; Length 469;
Best Local Similarity 86.7%; Pred. No. 9.8e-148;
Matches 364; Conservative 22; Mismatches 31; Indels 3; Gaps 1;
Qy 1 MATSGTVPFVPMVLCVCSAVSHRNQQTWFEVFLSSMCPINVSASTLYGIMFDAGSTGT 60
Db 1 MATPGAVFVFLMIAACAGSTVFYRQQTWFEVFLSSMCPANVSASTLYGIMFDAGSTGT 60
Qy 61 RIHVYTFVOKMPQLPILGEVDSVKPGLSAPVDOPKOGAETVQGLLEVAKDSIPRSHW 120
Db 61 RIHVYTFVOKAAGQLPFLGEIEFDSVKPGLSAPVDOPKOGAETVQGLLEVAKDSIPRSHW 120
Qy 121 KKTVPVVKATAGLRLLEPHKAKALLFEVKEIFRKSPLFPKGSVIMDGSDEGILAWTV 180
Db 121 KKTVPVVKATAGLRLLEPHKAKALLFEVKEIFRKSPLFPKGSVIMDGSDEGILAWTV 180
Qy 181 NFLTGQHGHRQBTGTDLGGASTQITFLPQPEKLEOTPRGYLTSFEMFNSTFKLYTH 240
Db 181 NFLTGQHGHRQBTGTDLGGASTQITFLPQPEKLEOTPRGYLTSFEMFNSTFKLYTH 240
Qy 241 SYLGFGKKAARLALATLGALETGTGHTFRSACLPRWLEAEWIFGGVKYQVGGNQEVEGF 300
Db 241 SYLGFGKKAARLALATLGALETGTGHTFRSACLPRWLEAEWIFGGVKYQVGGNQEVEGF 300
Qy 301 EPCYAEVLRVVRGKLHQPVEVQSGVAFYSYVYDRAVDTMDIDYKGGILKVEDFERKAR 360
Db 301 EPCYAEVLRVVRGKLHQPVEVQSGVAFYSYVYDRAVDTMDIDYKGGILKVEDFERKAR 360
Qy 361 EVCNLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQTKKNNIETGWALGATFHL 420
Db 361 EVCNLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQTKKNNIETGWALGATFHL 417

RESULT 4
ENP6_HUMAN
ID ENP6_HUMAN STANDARD; PRT; 484 AA.
AC Q75354; Q9JUD1;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Ectonucleoside triphosphate diphosphohydrolase 6 (EC 3.6.1.6)
DE (NTPase6) (CD39 antigen-like 2).
GN ENTPD6 OR CD39L2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98341119; PubMed=9676430;
RA Chawick B.P., Frischauf A.-M.;
RT "The CD39-like gene family: identification of three new human members
RT (CD39L2, CD39L3, and CD39L4), their murine homologues, and a member of

```





GN GDAl OR YEL042W OR SYGP-ORF16.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
RC STRAIN=G2-9;  
RX MEDLINE=93308137; PubMed=8391537;  
RA Abeijon C., Yangisawa K., Mandon E.C., Haeusler A., Moremen K.,  
RA Hirschberg C.B., Robbins P.W.;  
RT "Guanosine diphosphate is required for protein and sphingolipid  
glycosylation in the Golgi lumen of Saccharomyces cerevisiae.";  
RL J. Cell Biol. 122:307-323(1993).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=S288C / AB972;  
RX PubMed=9169868;  
RA Dietrich F.S., Mulligan J.T., Hennessey K.M., Yelton M.A., Allen E.,  
RA Araujo R., Aviles E., Berno A., Brennan T., Carpenter J., Chen E.,  
RA Cherry J.M., Chung E., Duncan M., Guzman E., Hartzell G.,  
RA Hunnicke-Smith S., Hyman R.W., Kayser A., Komp C., Lashkari D., Lew H.,  
RA Lin D., Mosedale D., Nakahara K., Namath A., Norgren R., Oefner P.,  
RA Oh C., Petel F.X., Roberts D., Sehl P., Schramm S., Shogren T.,  
RA Smith V., Taylor P., Wei Y., Botstein D., Davis R.W.;  
RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome V.";  
RL Nature 387:78-81(1997).  
CC -1- FUNCTION: AFTER TRANSFER OF SUGARS TO ENDOGENOUS MACROMOLECULAR  
ACCEPTORS, THE ENZYME CONVERTS NUCLEOSIDE DIPHOSPHATES TO  
NUCLEOSIDE MONOPHOSPHATES WHICH IN TURN EXIT THE GOLGI LUMEN IN  
A COUPLED ANTIPORTER REACTION, ALLOWING ENTRY OF ADDITIONAL  
NUCLEOTIDE SUGAR FROM THE CYTOSOL.  
CC -1- CATALYTIC ACTIVITY: GDP + H(2)O = GMP + phosphate.  
CC -1- PATHWAY: Glycosylation.  
CC -1- SUBCELLULAR LOCATION: Type II membrane protein. Golgi.  
CC -1- SIMILARITY: BELONGS TO THE GDAl / CD39 NTPASE FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; L19560; AAA34656.1; -;  
CC EMBL; U18779; AAB65000.1; -;  
CC PIR; A40732; A40732.  
CC SGD; S0000768; GDAl.  
CC GO; GO:0004382; F:guanosine diphosphatase activity; IDA.  
CC GO; GO:0045134; F:uridine diphosphatase activity; IDA.  
CC InterPro; IPR000407; GDAl\_CD39\_NTPase.  
CC Pfam; PF01150; GDAl\_CD39; 1.  
CC PROSITE; PS01238; GDAl\_CD39\_NTPASE; 1.  
CC HydroLase; Golgi stack; Glycoprotein; Transmembrane; Signal-anchor.  
FT DOMAIN 1 9  
FT TRANSMEM 10 24  
FT SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)  
FT (POTENTIAL).  
FT LUMENAL (POTENTIAL).  
FT CARBOHYD 25 518  
FT CARBOHYD 41 41 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 280 280 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 335 335 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 518 AA; 56821 MW; 9A61182D5ED22ADF CRC64;  
Query Match 21.5%; Score 483.5; DB 1; Length 518;  
Best Local Similarity 32.8%; Pred. No. 3.3e-31;  
Matches 142; Conservative 56; Mismatches 160; Indels 75; Gaps 15;  
49 YGIMFDAGSTGTRTHVYTF-VQKMPGQLPILGEGVDSVKPGLSAFVDQPKQGAETVQGL 107  
93 YVIMIDAGSTGSRVHIYKFDVCTSP---PTLLDEKFMLEFGLSFTDVGANSIDPL 149  
108 LEVAKDSIPRSHWKKTPVVLKATAGRLPPEHAKALLFEVKEIFRKS-PFLVPKG-SVS 165

Db 150 LKVMNVPIKARCTPVAVKATAGRLLLGDAKSSKILSAVRDHLKDYFPVVEGDVS 209  
Qy 166 IMGDSDEGILAWTVNFTLQGL--HGHRQETVCTDLDGGASTOITFLPQPEKTELEQTPRG 223  
Db 210 IMGDDEGVFAWITTVYLLGNIGANGPKLPTAAVFDLGGSTQIVFEPTFINEKWDGE 269  
Qy 224 YLTSFEMFNSTYKLYTHSYLGFGLKAAR-----LATLGALETEGTDGHTFSAC 272  
Db 270 HKFDLKFGDENYTLQFSLGYGLKGRNKNVSVLVENALKDGKILKGDNTKTHQLSSPC 329  
Qy 273 LPRWLEA--EWI-----FGVKVQYGG-----NOGEVGFEPFCYAEVLR 309  
Db 330 LPPKVATNEKVTLESKETVTFIDFGDEPSGAQCRFLTDEILNKAQACQSPSPF---- 385  
Qy 310 VVRGKLHQPVEV---QRGSFYAFSYVDRAVDTDM-IDYKGGILKVEDEPKAREVC- 363  
Db 386 ---NGVHQPSLVTRTFKESNDIYFYDTRPLGMLPSF-----TLNENLDLARLVCK 436  
Qy 364 -----DNLENFTSGSPFLCMDSYITALLKDGFGFADSTVLQLTKKVNNIET 410  
Db 437 GEETWNSVPSGIAGSLDELESDFH-CLDLSFOVSLHTGYDIPLOQLRATGKKIANKEI 495  
Qy 411 GWALGATFHLLOS 423  
Db 496 GWCLGASLPLKA 508

RESULT 8  
APY SOLTU  
ID AC APY SOLTU STANDARD; PRT; 454 AA.  
RC P80595; Q43164;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE Apyrase precursor (EC 3.6.1.5) (ATP-diphosphatase) (Adenosine  
diphosphatase) (ADPase) (ATP-diphosphohydrolase).  
GN RROPI.  
OS Solanum tuberosum (Potato).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
OC Asteridae; lamiales; Solanales; Solanaceae; Solanum.  
OX NCBI\_TaxID=4113;  
RN [1]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 59-160; 236-253 AND 332-345.  
RC TISSUE=Tube;  
RX MEDLINE=96158985; PubMed=8579614;  
RA Handa M., Guidotti G.;  
RT "Purification and cloning of a soluble ATP-diphosphohydrolase  
(apyrase) from potato tubers (Solanum tuberosum).";  
RL Biochem. Biophys. Res. Commun. 218:916-923(1996).  
RN [2]  
RP SEQUENCE OF 42-54; 68-95 AND 236-253.  
RC STRAIN=cv. Desiree;  
RX MEDLINE=96355615; PubMed=8703025;  
RA Vasconcelos E.G., Ferreira S.T., de Carvalho T.M.U., de Souza W.,  
RA Kettlun A.M., Mancilla M., Valenzuela M.A., Vertovski-Almeida S.;  
RT "Partial purification and immunohistochemical localization of ATP  
diphosphohydrolase from Schistosoma mansoni. Immunological cross-  
reactivities with potato apyrase and toxoplasma gondii nucleoside  
triphosphate hydrolase.";  
RL J. Biol. Chem. 271:22139-22145(1996).  
CC -1- FUNCTION: CATALYZES THE HYDROLYSIS OF PHOSPHOANHYDRIDE BONDS OF  
NUCLEOSIDE TRI- AND DI-PHOSPHATES.  
CC -1- CATALYTIC ACTIVITY: ATP + 2 H(2)O = AMP + 2 phosphate.  
CC -1- COFACTOR: CALCIUM.  
CC -1- SUBCELLULAR LOCATION: Membrane-associated (Probable).  
CC -1- PTM: THE N-TERMINUS IS BLOCKED.  
CC -1- SIMILARITY: BELONGS TO THE GDAl / CD39 NTPASE FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL: U58597; AAB02720.1; --  
 PIR: JC4616; JC4616.  
 InterPro: IPR000407; GDAI\_CD39\_NTPase.  
 Pfam: PF01150; GDAI\_CD39.1.  
 PROSITE: PS01238; GDAI\_CD39\_NTPase; 1.  
 Hydrolase; Transmembrane; Calcium; Signal.  
 SIGNAL 1 30 POTENTIAL.  
 CHAIN 31 454 APOLASE.  
 TRANSMEM 426 446 POTENTIAL.  
 CARBOHYD 151 151 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 CARBOHYD 262 262 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SEQUENCE 454 AA; 50041 MW; 9D9FE431DA2F52F CRC64;

Query Match 19.8%; Score 446.5; DB 1; Length 454;  
 Best Local Similarity 28.8%; Pred. No. 2.6e-28;  
 Matches 131; Conservative 78; Mismatches 177; Indels 69; Gaps 16;

QY 25 NQOTWEGIFLS-----SMCPINVA-----STLYGIMEDAGSGTGRHVV 65  
 DB 3 NQNSHFIFILALFLVPLSLSKNVNAQIPLRHLLSHSEHYAIFDAGSTGSRVRF 62  
 QY 66 TFQKMPGQPILEG-EVFDVSKPGLSAFADQPKQGAETVQGLLEVAKDSIPRSHWKTP 124  
 DB 63 RDEKL-GLLPIGNIEYFWATEPGLSSVAEDPKAANSLEPLDGAEGVVPQELQSETP 121  
 QY 125 VVKATAGRLRLPEHAKALLFEVKEIFR-KSPFLPKGSVSMDSGSGILAWVTNVL 183  
 DB 122 LELGATAGRLMLGDAAEKILQAVRLNVKNQSTFHSKDQWVTILDGTQBSYMAAINYL 181  
 QY 184 TQGLHGRQETVCTLDLGGASTQITEL---POPEKLE-QTPRGYLTSEMFNSTYKLYT 239  
 DB 182 LGNLGKDYKSTTATIDLGGSQVMAYAINQEPKAPQNEDEGPPYQOQKHLMSKDNLYV 241  
 QY 240 HSYLFGGLKAARLATLGALETECTDGTFRSACLPRWLAEMTFGGVQYQYGNQGEVVG 299  
 DB 242 HSYLNGQLAGRAEIEFKASRNES-----NPCALEGCDGYSGYGVQVYKAPKGS-- 292  
 QY 300 FFCYAEVLVRVG--KLHQPEEVQSGP-----YAFSYYYDRAVDTMI 342  
 DB 293 ---SWKRCRLTRHALKINAKCNIECTFNGVWNGGGQKQNIHASSPFYDGAQGVIV 349  
 QY 343 DYE-KGILKVEDEFAKREVCN-LENFTS-----GSPFLCMDLSYITALLKDGFG 392  
 DB 350 DTKFPSALAKPIQYLNAAKVACQTNVADTKSIFPKTQDRNIPYLCMDLIYEYTLVDGFG 409  
 QY 393 F---ADSTVLQLTQKVN--IETGWLGCATPHLLQS 423  
 DB 410 LNPKEITVIHDVQYKNYLVGAAPLGCALDVSS 444

## RESULT 9

YV4E\_CAEEL STANDARD; PRT; 485 AA.  
 AC Q18411;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Hypothetical 54.3 kDa protein C33H5.14 in chromosome IV.  
 GN C33H5.14.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;  
 RA Bradshaw H., Stellies L.;  
 RL Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.

## CC -!- SIMILARITY: BELONGS TO THE GDAI / CD39 NTPASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL: U41007; AAA82272.1; --  
 PIR: T34147; T34147.  
 WormPep: C33H5.14; CE04157.  
 InterPro: IPR000407; GDAI\_CD39\_NTPase.  
 Pfam: PF01150; GDAI\_CD39.1.  
 PROSITE: PS01238; GDAI\_CD39\_NTPase; 1.  
 Hypothetical protein; Transmembrane; Hydrolase.  
 TRANSMEM 439 459 POTENTIAL.  
 SEQUENCE 485 AA; 54309 MW; DE64DIADC20F581E CRC64;

Query Match 17.2%; Score 387.5; DB 1; Length 485;  
 Best Local Similarity 27.7%; Pred. No. 1.6e-23;  
 Matches 120; Conservative 84; Mismatches 152; Indels 77; Gaps 18;

QY 49 YGIMFDAGSTGRHVVTFVQKMPGQPILEGVFD-----VKPGLSAFVDOPKQGA 102  
 DB 25 YGVICDAGSGTGRFVYT-LKPLSGGLTNIDTLIHSESPVVKVTPGLSSFGDKPEQVVE 83  
 QY 103 TVQGLEVAKDSIPRSHWKTPVVLKATAGRLRLPEHAKALLFEVKEIFRSPFL-VPK 161  
 DB 84 YUTPLRLFAEHIPIEQLGETDILLFATAGMRLLEPAQKDAIKNLQGLKSVTALRVSD 143  
 QY 162 GSVSMDSGDEGILAWVTNFLTGLHGRQETVCTLDLGGASTQITFLPQPEK----- 215  
 DB 144 SNRIIDGAWEGISWIAVNYILGRFDKENDSKVGMIDMGASVQIAFBIANEKESYNG 203  
 QY 216 TLEQTPRGVLTSEMFNSTYKLYTHSYLGFGLKAARLATLGALETECTDGTFRSACLPR 275  
 DB 204 NYVEINLGSITNEDYK--YKYSTTFLGYGANEGLKRYENSLVKSGNS-----NDSCSPR 257  
 QY 276 WLEAEWIFGGVQYQYGNQGEVGPCEVAVLVRVVRGKLHOPE----- 319  
 DB 258 GLNR--LIG-----EFTVNGTGE--WDVCLAQVSSLI-GDKAQPSCNPFCFLRNVTAPSV 308  
 QY 320 EYQGSFYAFS--YYDRAVDTMDIDYKGGILKVEDFERKAREVC-----DNLENFTSGSP 374  
 DB 309 NLSTVQLYGFSEYWTTS-----NFGSGGEYHYQKFTDEVRYKQCKDNDIQDGFKRNE 362  
 QY 375 F-----LCMDLSYITALLKDGFGPADST--VLQTKKVNNIETGWLGA----- 416  
 DB 363 FPNADIERLGTNCFKAAWVTSVLHDGFN-VDKTKHLFQSVLKIAGEMQWALGAMLYHSK 421  
 QY 417 --TFHLLQSLGIS 427  
 DB 422 DLKFNLLQLEVA 434

## RESULT 10

YBU4\_CAEEL STANDARD; PRT; 552 AA.  
 ID YBU4\_CAEEL  
 AC Q21815;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DE Hypothetical protein R07E4.4 in chromosome X.  
 GN R07E4.4.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Bristol N2;



RA Miller N.:  
 RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP REVISIONS.  
 RA Waterston R.;  
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: BELONGS TO THE GDAL / CD39 NTPASE FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC  
 CC EMBL; U39652; AAA80403.2; -  
 CC WormPep; R07E4.4; CE28748.  
 DR InterPro; IPR000407; GDAL\_CD39\_NTPase.  
 DR Pfam; PF01150; GDAL\_CD39; 1.  
 DR PROSITE; PS01238; GDAL\_CD39\_NTPASE; 1.  
 KW Hypothetical protein; Transmembrane; Hydrolase.  
 FT TRANSMEM 7 27 POTENTIAL.  
 FT TRANSMEM 490 510 POTENTIAL.  
 SQ SEQUENCE 552 AA; 62510 MW; E5DC32C858AE4D94 CRC64;  
 Query Match 17.1%; Score 385; DB 1; Length 552;  
 Best Local Similarity 28.6%; Pred. No. 3e-23;  
 Matches 126; Conservative 69; Mismatches 160; Indels 86; Gaps 20;  
 QY 49 YGIMFDAGSTGTRIHYVTFVQKPGQPLILEGEVDFS-----VKPLSAFVDPQKQAE 102  
 DB 44 YGVICDAGSTGTRLFVYNWISTSDSELIQIEPVIYDNKPMVKISPLGLSTFGTKPAQAE 103  
 QY 103 TVGGLLEVAKDSIPRSHWKTPTVVLKATAGRLRLPEHKAKALLFEVKEIERKSPFL----- 158  
 DB 104 YLRPLMELAEHRHPEEKRPYTPVFIATAGRLRLPDEQKEAVL---KNLRNKLPKITSMQ 160  
 QY 159 VPKGSVIMDSDEGILAWTVNPLTQLH-----GH-RQETVGTLDLGGAST 205  
 DB 161 VLKEHRIIEGKEGIVSYIAVAVLAKGNKNTATLDPFGSPAHARQKTVMIDMGASA 220  
 QY 206 QITF-LPQPK-----TLEQTPRGVLTSFENPNSTYKLYTHSYLGF-----GLKARLALIG 256  
 DB 221 QIAFELPDTDFSSINVENINLGRDSDSLFK--YKLFVTTLFLGYGVNEGIRKYEMLLS 278  
 QY 257 ALTEGTDGHTFSACLPRLEAEWIFGGVQYQGN--QEG-----E 297  
 DB 279 KLKQD--NGVIQDCMPLNLHKT-----VTLNGENFVRRTGNWNTCSNEVKKLLNPE 331  
 QY 298 VGFEPCEVLRVVRGKLHOPE-EVORGSPYAFSYDDRAVDMDIDYKGGILKVEDFE 356  
 DB 332 SSSEVCKAEAAKCYGAVPAPSPISLNIENYGFSEYWYTHDV----LGLGGQYDAENIA 387  
 QY 357 RKAREVCD-----NLENFTSGSP-----FLCMDLSVITALLKDGFGFADST--VLQ 400  
 DB 388 KKTQYCKSKWSTQAEKSKQLYPRADEERLRTQCFSKAWITSLVLDGFS-VDKTHNFKQ 446  
 QY 401 LTKKVNNIETGVALGA--TEHL 420  
 DB 447 SVSTIAGQEQVQWALGAMIYHM 467  
 RESULT 11  
 ID\_ENP1\_CHICK STANDARD; PRT; 493 AA.  
 AC Q93295;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Ectonucleoside triphosphate diphosphohydrolase 1 (EC 3.6.1.5)  
 DE (NTPase) (Ecto-ATP diphosphohydrolase) (ATPase) (lymphoid cell  
 DE activation antigen) (Ecto-apyrase) (CD39 antigen).  
 GN ENTPD1 OR CD39.  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OX NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-21 AND 150-156.  
 RC TISSUE=Oviduct;  
 RX MEDLINE=98298108; PubMed=9632655;  
 RA Nagy A.K., Knowles A.F., Nagami G.T.;  
 RT "Molecular cloning of the chicken oviduct ecto-ATP-  
 RT diphosphohydrolase.";  
 RL J. Biol. Chem. 273:16043-16049(1998).  
 RN [2]  
 RP SEQUENCE OF 1-17.  
 RC TISSUE=Stomach;  
 RX MEDLINE=97442428; PubMed=9295305;  
 RA Lewis-Carl S., Kirley T.L.;  
 RT "Immunolocalization of the ecto-ATPase and ecto-apyrase in chicken  
 RT gizzard and stomach. Purification and N-terminal sequence of the  
 RT stomach ecto-apyrase.";  
 RL J. Biol. Chem. 272:23645-23652(1997).  
 CC -1- FUNCTION: In the nervous system, could hydrolyze ATP and other  
 CC nucleotides to regulate purinergic neurotransmission. Could also  
 CC be implicated in the prevention of platelet aggregation.  
 CC Hydrolyzes ATP and ADP equally well (by similarity).  
 CC -1- CATALYTIC ACTIVITY: ATP + 2 H(2)O = AMP + 2 phosphate.  
 CC -1- COFACTOR: REQUIRES CALCIUM AND MAGNESIUM (BY SIMILARITY).  
 CC -1- SUBUNIT: HOMODIMER; DISULFIDE-LINKED (PROBABLE).  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
 CC -1- PTM: N-GLYCOSYLATED.  
 CC -1- SIMILARITY: BELONGS TO THE GDAL / CD39 NTPASE FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC  
 CC EMBL; AF041355; AAC26491.1; -  
 DR InterPro; IPR000407; GDAL\_CD39\_NTPase.  
 DR Pfam; PF01150; GDAL\_CD39; 1.  
 DR PROSITE; PS01238; GDAL\_CD39\_NTPASE; 1.  
 KW Hydrolase; Transmembrane; Antigen; Glycoprotein; Calcium; Magnesium.  
 FT DOMAIN 1 7 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 8 28 POTENTIAL.  
 FT DOMAIN 29 463 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 464 486 POTENTIAL.  
 FT DOMAIN 487 493 CYTOPLASMIC (POTENTIAL).  
 FT CARBOHYD 65 65 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 79 79 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 133 133 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 223 223 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 234 234 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 267 267 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 324 324 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 330 330 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 361 361 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 372 372 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 382 382 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 445 445 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CONFLICT 16 16 C -> W (IN REF. 2).  
 FT CONFLICT 21 21 I -> G (IN REF. 1; AA SEQUENCE).  
 SQ SEQUENCE 493 AA; 54034 MW; F14FF4C3AA2F3603 CRC64;  
 Query Match 17.0%; Score 383; DB 1; Length 493;  
 Best Local Similarity 27.3%; Pred. No. 3.7e-23;  
 Matches 131; Conservative 66; Mismatches 170; Indels 112; Gaps 18;  
 QY 11 MLVWSCVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMPDAAGTGTGRIHYTF-VQ 69



Db 11 LTATCVFIIILILSAVDKDFL-----PPGTYKGLVFDAGSTHTALVYVQMPAD 62  
 QY 70 KMPGQLPILEGEVDFSVKPGLSAFVQPKQGAETVQGLLEVAKDSIPRSHWKKTPVVLKA 129  
 Db 63 KENGTVQVSECTVNGSGISSYADDAGAGASLXPCLDKAMAVIPVQOWCTPYLGA 122  
 QY 130 TAGRLRLPEH---KAKALIFEVKEIFRKGFPFLVPGKSVIMDSDEGLILAWTVNFLTQ 186  
 Db 123 TAGMRLLEQNSTKAEQVFAEVSKAIRFP--VDFRGAQILTNGEESFGMITVNYLLET 180  
 QY 187 L-----HCHROETVGLDLCAGASTQITLPLQFEKTELTQTPRGYLTSPFMFNSTYK 236  
 Db 181 LTKFSAGKWEHPQNTVEVLGALDLCAGASTQITPQGV--TIEDKNTSVL--FRLYGTNTYS 236  
 QY 237 LYTHSVLGFGLKARLALCALTEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGNQOE- 295  
 Db 237 LYTHSVLVCQGOQASKRLMAALHQDSYVQNISHPCYK-----GYRLI 280  
 QY 296 ---GEVGFPCY-----AEVLVR-----GKLHQ 317  
 Db 281 ITIAEYDPCVPTPMSLSPAQLITVTGTGNPAACPTAILKLFNLTCGANRTCGPDGVYQ 340  
 QY 318 PEEVQGSFYAFS-YYDRA-----VDTMDIY-EKGGILKVEDFERKAREVC 363  
 Db 341 PP--VRGQFFAFAGFYTTFSFLNLTGQSLSHVNATVDFCNKNWSELVETFPQNK---- 394  
 QY 364 DNLENFTSGSPFPCMDLSYITALLKDGFGFADST--VLQLTQKVNNIETGALGATFHL 420  
 Db 395 EHLHTY-----CVVGLYILTLVDGYKFEHTWNSNIHFSQAKNAGNIGWTLGFMNLN 446

RESULT 12  
 YND1\_YEAST STANDARD; PRT; 630 AA.  
 AC F40009; 1995 (Rel. 31, Created)  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Golgi apyrase (EC 3.6.1.5) (ATP-diphosphatase) (Adenosine  
 DE diphosphatase) (ADPase) (ATP-diphosphohydrolase) (Golgi nucleoside  
 DE diphosphatase).  
 GN YND1 OR YER005W.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND CHARACTERIZATION.  
 RC STRAIN=S288C;  
 RX MEDLINE=99340091; PubMed=10409709;  
 RA Gao X.D., Kaigorodov V., Jigami Y.,  
 RT "YND1, a homologue of GDAL, encodes membrane-bound apyrase required  
 RT for Golgi N- and O-glycosylation in Saccharomyces cerevisiae.";  
 RL J. Biol. Chem. 274:21450-21456(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=S288C / AB972;  
 RX PubMed=9169868;  
 RA Dietrich F.S., Mulligan J.T., Hennessy K.M., Yelton M.A., Allen E.,  
 RA Araujo R., Aviles E., Berno A., Brennan T., Carpenter J., Chen E.,  
 RA Cherry J.M., Chung E., Duncan M., Guzman E., Hartzell G.,  
 RA Hunnicke-Smith S., Hyman R.W., Kayser A., Komp C., Lashkari D., Lew H.,  
 RA Lin D., Mosedale D., Nakahara K., Namath A., Norgren R., Oefner P.,  
 RA On C., Petel F.X., Roberts D., Sehl P., Schramm S., Shogren T.,  
 RA Smith V., Taylor P., Wei Y., Botstein D., Davis R.W.;  
 RT "The nucleotide sequence of Saccharomyces cerevisiae chromosome V.";  
 RL Nature 387:78-81(1997).  
 CC -1- FUNCTION: CATALYZES THE HYDROLYSIS OF PHOSPHOANHYDRIDE BONDS OF  
 CC NUCLEOSIDE TRI- AND DI-PHOSPHATES. HAS EQUAL HIGH ACTIVITY TOWARD  
 CC ADP/ATP, GDP/GTP, AND UDP/UTP AND APPROXIMATELY 50% LESS TOWARD  
 CC CDP/CTP AND THIAMINE PYROPHOSPHATE. HAS NO ACTIVITY TOWARD GMP.  
 CC REQUIRED FOR GOLGI GLYCOSYLATION AND CELL WALL INTEGRITY.

CC -1- CATALYTIC ACTIVITY: ATP + 2 H(2)O = AMP + 2 phosphate.  
 CC -1- PATHWAY: Glycosylation.  
 CC -1- SUBCELLULAR LOCATION: MEMBRANE BOUND; GOLGI.  
 CC -1- SIMILARITY: BELONGS TO THE GDAL / CD39 NTPASE FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC  
 CC EMBL; AF203695; AAF17573.1; -.  
 DR EMBL; U18778; AAB64538.1; -.  
 DR PIR; S50463; S50463.  
 DR SGD; S000807; YND1.  
 DR GO; GO:0006486; P:protein amino acid glycosylation; IMP.  
 DR InterPro; IPR000407; GDAL\_CD39\_NTPase.  
 DR Pfam; PF01150; GDAL\_CD39; 1.  
 DR PROSITE; PS01238; GDAL\_CD39\_NTPASE; 1.  
 KW Hydrolase; Transmembrane; Golgi stack.  
 FT DOMAIN 1 500 LUMENAL (POTENTIAL).  
 FT TRANSMEM 501 517 POTENTIAL.  
 FT DOMAIN 518 630 CYTOPLASMIC (POTENTIAL).  
 FT SEQUENCE 630 AA; 71851 MW; 02F8D24A78212544 CRC64;  
 SQ  
 Query Match 16.5%; Score 372; DB 1; Length 630;  
 Best Local Similarity 26.2%; Pred. No. 3.9e-22;  
 Matches 117; Conservative 73; Mismatches 165; Indels 92; Gaps 17;  
 QY 49 YGIMFDAGSTGRIHYTF-----VQKMPGQLPILEGE--VFDSVKPGLSA 92  
 Db 10 FGVIVDAGSSGSRHVFQKQWDTESLLHATNQDSQSLQSVPHIQHOKDWTFF-KLNFGLS 68  
 QY 93 FVDQPKQGAET-VQGLLEVAKDSIPRSHWKKTPVVLKATAGRLRLPEHAKALLFEV-KE 150  
 Db 69 FEKKPDQAYKSHIKPLLPFAKNIIPESHWSCEPVFIQATAGMRLLPQDIQSSILDLGCG 128  
 QY 151 IFRKSPFLVP--KGSVSIWDGSDGLVAVTVNFLTQGLHGRQE-----TVGTLDLGA 203  
 Db 129 LKHPAEFLVEDCSAQIQVIDGETEGYLGWGLNLYLGHFNNDYNPEVSDHDTFGFMDMGA 188  
 QY 204 STQITFLPQFEKLE-----QTPRGYLTSEFMFNSTYKLYTHSVLGFGLKAARL 252  
 Db 189 STQIAFAPHDSGEIARHRDIIATIFLRSVNGDLQKWDVFTST-----WLGFGANQARR 241  
 QY 253 ATLGAL-----ETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGNQOEGEVGPE 301  
 Db 242 RYLAQLINTLPENTNDYENDDFSTRNLNDPCMPRGSSDTDFPKDTIFHAGSGNYEQCTK 301  
 QY 302 PCVAEVLK-----VVRGKLHOP-EVQSGSFYAFSVYVYDRAVDTMDIDYKGGILKV 352  
 Db 302 SYPLPLKKNPCDDPECLFNGVHAPRIDPANKFICTSEYVYVYANDV-----FKLGGEYNF 357  
 QY 353 EDFERKAREVCN-----LENFTSG-----SPFL---CMDLSYITALLKDGFGFA---- 394  
 Db 358 DKFESLSREFCNSNWTQILANSKGVYNSIPENFLKDACFKGNWNLNHLHGFDMPRIDV 417  
 QY 395 -----DSTVLQTLTKKVNNIETGALG 415  
 Db 418 DAENVNDRPLFQSVKEKVERELSWTLG 444  
 RESULT 13  
 ENP1\_MOUSE STANDARD; PRT; 510 AA.  
 AC P55772;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Ectonucleoside triphosphate diphosphohydrolase 1 (EC 3.6.1.5)  
 DE (NTPDase1) (Ecto-ATP diphosphohydrolase) (ATPDase) (Lymphoid cell







GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 10:05:35 ; Search time 29.5 Seconds  
(without alignments)  
3015.210 Million cell updates/sec

Title: US-10-091-085-7  
Perfect score: 2252  
Sequence: 1 MATSWGTVMFLVVCVCSA.....ETCWALGATFLLQLSLGISH 428

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 789580 seqs, 207824079 residues  
Total number of hits satisfying chosen parameters: 789580

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2252	100.0	428	12	US-10-286-926-7
2	2252	100.0	428	14	US-10-091-085-7
3	2252	100.0	428	14	US-10-092-063-7
4	2235	99.2	428	12	US-10-286-926-3
5	2235	99.2	428	12	US-10-286-926-5
6	2235	99.2	428	12	US-10-231-913-127
7	2235	99.2	428	14	US-10-091-085-3
8	2235	99.2	428	14	US-10-091-085-5
9	2235	99.2	428	14	US-10-092-063-3
10	2235	99.2	428	14	US-10-092-063-5
11	2089	92.8	405	12	US-10-286-926-25
12	2089	92.8	405	14	US-10-092-063-25
13	1979.5	87.9	427	12	US-10-231-913-126
14	1822.5	80.9	465	14	US-10-092-063-39
15	1645	73.0	330	9	US-09-925-299-876

16	1645	73.0	330	11	US-09-925-299-876	Sequence 876, App
17	996	44.2	484	12	US-10-286-926-27	Sequence 27, Appl
18	996	44.2	484	12	US-10-231-913-123	Sequence 123, App
19	996	44.2	484	14	US-10-092-063-27	Sequence 27, Appl
20	993	44.1	467	12	US-10-231-913-36	Sequence 36, Appl
21	990.5	44.0	379	12	US-10-231-913-271	Sequence 271, App
22	989	43.9	484	12	US-10-231-913-124	Sequence 124, App
23	986	43.8	446	12	US-10-231-913-38	Sequence 38, Appl
24	983	43.7	455	12	US-10-231-913-125	Sequence 125, App
25	616.5	27.4	479	12	US-10-369-493-6447	Sequence 6447, Ap
26	605	26.9	476	9	US-09-835-147-4	Sequence 4, Appli
27	548	24.3	476	9	US-09-835-147-3	Sequence 3, Appli
28	535.5	23.8	556	12	US-10-369-493-2169	Sequence 2169, Ap
29	507.5	22.5	467	9	US-09-129-113-19	Sequence 19, Appl
30	502.5	22.3	459	9	US-09-129-113-9	Sequence 9, Appl
31	498	22.1	410	12	US-10-231-913-272	Sequence 272, App
32	493	21.9	462	9	US-09-129-113-2	Sequence 2, Appli
33	488.5	21.7	518	12	US-10-369-493-1713	Sequence 1713, Ap
34	486.5	21.6	462	9	US-09-129-113-15	Sequence 15, Appl
35	432	19.2	455	12	US-10-259-165-286	Sequence 286, App
36	426.5	18.9	459	12	US-10-259-165-74	Sequence 74, Appl
37	422	18.7	605	12	US-10-369-493-3333	Sequence 3333, Ap
38	382.5	17.0	485	12	US-10-369-493-6055	Sequence 6055, Ap
39	380	16.9	495	9	US-09-823-356-4	Sequence 4, Appli
40	374.5	16.6	557	12	US-10-369-493-6789	Sequence 6789, Ap
41	371.5	16.5	458	12	US-10-199-672-496	Sequence 496, App
42	371.5	16.5	458	12	US-10-187-749-496	Sequence 496, App
43	371.5	16.5	458	12	US-10-194-457-496	Sequence 496, App
44	371.5	16.5	458	12	US-10-184-642-496	Sequence 496, App
45	371.5	16.5	458	12	US-10-196-747-496	Sequence 496, App

ALIGNMENTS

RESULT 1  
US-10-286-926-7  
; Sequence 7, Application US/10286926  
; Publication No. US20030175752A1  
; GENERAL INFORMATION:  
; APPLICANT: Ford, John  
; APPLICANT: Mulero, Julio  
; APPLICANT: Yeung, George  
; TITLE OF INVENTION: Methods and Materials Relating to CD39-Like  
; TITLE OF INVENTION: Polypeptides  
; FILE REFERENCE: 28110/36457CON  
; CURRENT APPLICATION NUMBER: US/10/286,926  
; CURRENT FILING DATE: 2002-11-01  
; PRIOR APPLICATION NUMBER: 09/557,800  
; PRIOR FILING DATE: 2000-04-25  
; PRIOR APPLICATION NUMBER: 09/481,238  
; PRIOR FILING DATE: 2000-01-11  
; PRIOR APPLICATION NUMBER: 09/370,265  
; PRIOR FILING DATE: 1999-08-09  
; PRIOR APPLICATION NUMBER: PCT/US99/16180  
; PRIOR FILING DATE: 1999-07-16  
; PRIOR APPLICATION NUMBER: 09/350836  
; PRIOR FILING DATE: 1999-07-09  
; PRIOR APPLICATION NUMBER: 09/273447  
; PRIOR FILING DATE: 1999-03-19  
; PRIOR APPLICATION NUMBER: 09/122449  
; PRIOR FILING DATE: 1998-07-24  
; PRIOR APPLICATION NUMBER: 09/244444  
; PRIOR FILING DATE: 1999-02-04  
; PRIOR APPLICATION NUMBER: 09/118,205  
; PRIOR FILING DATE: 1998-07-16  
; NUMBER OF SEQ ID NOS: 54  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 7  
; LENGTH: 428  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-286-926-7

Query Match 100.0%; Score 2252; DB 12; Length 428;  
Best Local Similarity 100.0%; Pred. No. 1.6e-218;  
Matches 428; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MATSWGTFFMLVSVCSAVSHRNQOTWTFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60  
Db 1 MATSWGTFFMLVSVCSAVSHRNQOTWTFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

Qy 61 RIHVYTFVQKMPGOLPILGEVFDVSKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120  
Db 61 RIHVYTFVQKMPGOLPILGEVFDVSKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120

Qy 121 KKTTPVVLKATAGRLRLPEHKAALLFEVKEIFRKSPLVPKGSVIMTGDQEGIFAWTV 180  
Db 121 KKTTPVVLKATAGRLRLPEHKAALLFEVKEIFRKSPLVPKGSVIMTGDQEGIFAWTV 180

Qy 181 NFLTGQLHGRQETVGTLDLGASTQITFLPQPKETLEOTPRGYLTSFEMFNSTYKLYTH 240  
Db 181 NFLTGQLHGRQETVGTLDLGASTQITFLPQPKETLEOTPRGYLTSFEMFNSTYKLYTH 240

Qy 241 SYLGFGKAAARLATLGALETGDTGHTFRSACLPRWLEAEWIFGGVYQYGGNQGEGVGF 300  
Db 241 SYLGFGKAAARLATLGALETGDTGHTFRSACLPRWLEAEWIFGGVYQYGGNQGEGVGF 300

Qy 301 EPCYAEVLVRVVRGKLPQEEVQVGSFYAFSYYYDRAVDVTMDIDYEKGGILKVEDFERKAR 360  
Db 301 EPCYAEVLVRVVRGKLPQEEVQVGSFYAFSYYYDRAVDVTMDIDYEKGGILKVEDFERKAR 360

Qy 361 EVCNLENFTSGSPFLCMDLSYITALLKQGFADSTVLQTKKVNNIETGALGATFHL 420  
Db 361 EVCNLENFTSGSPFLCMDLSYITALLKQGFADSTVLQTKKVNNIETGALGATFHL 420

Qy 421 LQSLGISH 428  
Db 421 LQSLGISH 428

RESULT 2  
US-10-091-085-7  
; Sequence 7, Application US/10091085  
; Publication No. US20020146772A1  
; GENERAL INFORMATION:  
; APPLICANT: Ford, John  
; APPLICANT: Mulero, Julio  
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE POLYPEPTIDES  
; FILE REFERENCE: 28110/35761  
; CURRENT APPLICATION NUMBER: US/10/091.085  
; PRIOR FILING DATE: 2002-03-05  
; PRIOR APPLICATION NUMBER: 09/350,836  
; PRIOR FILING DATE: 1999-07-09  
; PRIOR APPLICATION NUMBER: 09/273,447  
; PRIOR FILING DATE: 1999-03-19  
; PRIOR APPLICATION NUMBER: 09/118,205  
; PRIOR FILING DATE: 1998-07-16  
; PRIOR APPLICATION NUMBER: 09/122,449  
; PRIOR FILING DATE: 1998-07-24  
; PRIOR APPLICATION NUMBER: 09/244,444  
; PRIOR FILING DATE: 1999-02-04  
; NUMBER OF SEQ ID NOS: 23  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 7  
; LENGTH: 428  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-091-085-7

Query Match 100.0%; Score 2252; DB 14; Length 428;  
Best Local Similarity 100.0%; Pred. No. 1.6e-218;  
Matches 428; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MATSWGTFFMLVSVCSAVSHRNQOTWTFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60  
Db 1 MATSWGTFFMLVSVCSAVSHRNQOTWTFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

Db 1 MATSWGTFFMLVSVCSAVSHRNQOTWTFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60  
Qy 61 RIHVYTFVQKMPGOLPILGEVFDVSKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120  
Db 61 RIHVYTFVQKMPGOLPILGEVFDVSKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120

Qy 121 KKTTPVVLKATAGRLRLPEHKAALLFEVKEIFRKSPLVPKGSVIMTGDQEGIFAWTV 180  
Db 121 KKTTPVVLKATAGRLRLPEHKAALLFEVKEIFRKSPLVPKGSVIMTGDQEGIFAWTV 180

Qy 181 NFLTGQLHGRQETVGTLDLGASTQITFLPQPKETLEOTPRGYLTSFEMFNSTYKLYTH 240  
Db 181 NFLTGQLHGRQETVGTLDLGASTQITFLPQPKETLEOTPRGYLTSFEMFNSTYKLYTH 240

Qy 241 SYLGFGKAAARLATLGALETGDTGHTFRSACLPRWLEAEWIFGGVYQYGGNQGEGVGF 300  
Db 241 SYLGFGKAAARLATLGALETGDTGHTFRSACLPRWLEAEWIFGGVYQYGGNQGEGVGF 300

Qy 301 EPCYAEVLVRVVRGKLPQEEVQVGSFYAFSYYYDRAVDVTMDIDYEKGGILKVEDFERKAR 360  
Db 301 EPCYAEVLVRVVRGKLPQEEVQVGSFYAFSYYYDRAVDVTMDIDYEKGGILKVEDFERKAR 360

Qy 361 EVCNLENFTSGSPFLCMDLSYITALLKQGFADSTVLQTKKVNNIETGALGATFHL 420  
Db 361 EVCNLENFTSGSPFLCMDLSYITALLKQGFADSTVLQTKKVNNIETGALGATFHL 420

Qy 421 LQSLGISH 428  
Db 421 LQSLGISH 428

RESULT 3  
US-10-092-063-7  
; Sequence 7, Application US/10092063  
; Publication No. US20020173005A1  
; GENERAL INFORMATION:  
; APPLICANT: Ford, John  
; APPLICANT: Mulero, Julio  
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE POLYPEPTIDES  
; FILE REFERENCE: 28110/35908  
; CURRENT APPLICATION NUMBER: US/10/092.063  
; CURRENT FILING DATE: 2002-03-05  
; PRIOR APPLICATION NUMBER: 09/370,265  
; PRIOR FILING DATE: 2002-01-31  
; PRIOR APPLICATION NUMBER: PCT/US99/16180  
; PRIOR FILING DATE: 1999-07-16  
; PRIOR APPLICATION NUMBER: 09/350,836  
; PRIOR FILING DATE: 1999-07-09  
; PRIOR APPLICATION NUMBER: 09/273,447  
; PRIOR FILING DATE: 1999-03-19  
; PRIOR APPLICATION NUMBER: 09/244,444  
; PRIOR FILING DATE: 1999-02-04  
; PRIOR APPLICATION NUMBER: 09/122,449  
; PRIOR FILING DATE: 1998-07-24  
; PRIOR APPLICATION NUMBER: 09/118,205  
; PRIOR FILING DATE: 1998-07-16  
; NUMBER OF SEQ ID NOS: 39  
; SOFTWARE: Patent In Ver. 2.0  
; SEQ ID NO 7  
; LENGTH: 428  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-092-063-7

Query Match 100.0%; Score 2252; DB 14; Length 428;  
Best Local Similarity 100.0%; Pred. No. 1.6e-218;  
Matches 428; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MATSWGTFFMLVSVCSAVSHRNQOTWTFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60  
Db 1 MATSWGTFFMLVSVCSAVSHRNQOTWTFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

61 RIHVYTFVQKMPGOLPILLEGVDFSVKPGLSAFVDPQKQGAETVQGLLEVAKDSIPRSHW 120  
61 RIHVYTFVQKMPGOLPILLEGVDFSVKPGLSAFVDPQKQGAETVQGLLEVAKDSIPRSHW 120  
121 KKTVPVLKATAGRLLLPEHKAALLFEVKEIFRKSPFLVPKGSVIMTQDDEGIFAWTV 180  
121 KKTVPVLKATAGRLLLPEHKAALLFEVKEIFRKSPFLVPKGSVIMTQDDEGIFAWTV 180  
181 NFLTQGLHGHROETVGTLDLGASTQITFLPOFEKTLQTPRGYLTSEMFNNTYKLYTH 240  
181 NFLTQGLHGHROETVGTLDLGASTQITFLPOFEKTLQTPRGYLTSEMFNNTYKLYTH 240  
241 SYLGFGLKAARLATIAGLETGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEVEGF 300  
241 SYLGFGLKAARLATIAGLETGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEVEGF 300  
301 EPCYAEVLVRVGRKLHQPVEVORGSGFYAFSYDDRAVDMDIDYKGGILKVEDFERKAR 360  
301 EPCYAEVLVRVGRKLHQPVEVORGSGFYAFSYDDRAVDMDIDYKGGILKVEDFERKAR 360  
361 EVCNLENFTSGSPFLCNDLSYITALLKDGFGFADSTVLQLTCKVNNIETGWALGATPHL 420  
361 EVCNLENFTSGSPFLCNDLSYITALLKDGFGFADSTVLQLTCKVNNIETGWALGATPHL 420  
421 LQSLGISH 428  
421 LQSLGISH 428

RESULT 4

US-10-286-926-3  
; Sequence 3, Application US/10286926  
; Publication No. US20030175752A1  
; GENERAL INFORMATION:  
; APPLICANT: Ford, John  
; APPLICANT: Mulero, Julio  
; APPLICANT: Yeung, George  
; TITLE OF INVENTION: Methods and Materials Relating to CD39-Like  
; FILE REFERENCE: 28110/36457CON  
; CURRENT APPLICATION NUMBER: US/10/286,926  
; CURRENT FILING DATE: 2002-11-01  
; PRIOR APPLICATION NUMBER: 09/557,800  
; PRIOR FILING DATE: 2000-04-25  
; PRIOR APPLICATION NUMBER: 09/481,238  
; PRIOR FILING DATE: 2000-01-11  
; PRIOR APPLICATION NUMBER: 09/370,265  
; PRIOR FILING DATE: 1999-08-09  
; PRIOR APPLICATION NUMBER: PCT/US99/16180  
; PRIOR FILING DATE: 1999-07-16  
; PRIOR APPLICATION NUMBER: 09/350836  
; PRIOR FILING DATE: 1999-03-19  
; PRIOR APPLICATION NUMBER: 09/273447  
; PRIOR FILING DATE: 1998-07-24  
; PRIOR APPLICATION NUMBER: 09/122449  
; PRIOR FILING DATE: 1998-02-04  
; PRIOR APPLICATION NUMBER: 09/118,205  
; PRIOR FILING DATE: 1998-07-16  
; NUMBER OF SEQ ID NOS: 54  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 428  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-286-926-3

Query Match 99.2%; Score 2235; DB 12; Length 428;  
Best Local Similarity 99.3%; Pred. No. 8.5e-217;  
Matches 425; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 1 MATSMGTFFMLVSCVCSAVSHRNQQTWFGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

1 MATSMGTFFMLVSCVCSAVSHRNQQTWFGIFLSSMCPINVSASTLYGIMFDAGSTGT 60  
61 RIHVYTFVQKMPGOLPILLEGVDFSVKPGLSAFVDPQKQGAETVQGLLEVAKDSIPRSHW 120  
61 RIHVYTFVQKMPGOLPILLEGVDFSVKPGLSAFVDPQKQGAETVQGLLEVAKDSIPRSHW 120  
121 KKTVPVLKATAGRLLLPEHKAALLFEVKEIFRKSPFLVPKGSVIMTQDDEGIFAWTV 180  
121 KKTVPVLKATAGRLLLPEHKAALLFEVKEIFRKSPFLVPKGSVIMTQDDEGIFAWTV 180  
181 NFLTQGLHGHROETVGTLDLGASTQITFLPOFEKTLQTPRGYLTSEMFNNTYKLYTH 240  
181 NFLTQGLHGHROETVGTLDLGASTQITFLPOFEKTLQTPRGYLTSEMFNNTYKLYTH 240  
241 SYLGFGLKAARLATIAGLETGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEVEGF 300  
241 SYLGFGLKAARLATIAGLETGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEVEGF 300  
301 EPCYAEVLVRVGRKLHQPVEVORGSGFYAFSYDDRAVDMDIDYKGGILKVEDFERKAR 360  
301 EPCYAEVLVRVGRKLHQPVEVORGSGFYAFSYDDRAVDMDIDYKGGILKVEDFERKAR 360  
361 EVCNLENFTSGSPFLCNDLSYITALLKDGFGFADSTVLQLTCKVNNIETGWALGATPHL 420  
361 EVCNLENFTSGSPFLCNDLSYITALLKDGFGFADSTVLQLTCKVNNIETGWALGATPHL 420  
421 LQSLGISH 428  
421 LQSLGISH 428

RESULT 5

US-10-286-926-5  
; Sequence 5, Application US/10286926  
; Publication No. US20030175752A1  
; GENERAL INFORMATION:  
; APPLICANT: Ford, John  
; APPLICANT: Mulero, Julio  
; APPLICANT: Yeung, George  
; TITLE OF INVENTION: Methods and Materials Relating to CD39-Like  
; FILE REFERENCE: 28110/36457CON  
; CURRENT APPLICATION NUMBER: US/10/286,926  
; CURRENT FILING DATE: 2002-11-01  
; PRIOR APPLICATION NUMBER: 09/557,800  
; PRIOR FILING DATE: 2000-04-25  
; PRIOR APPLICATION NUMBER: 09/481,238  
; PRIOR FILING DATE: 2000-01-11  
; PRIOR APPLICATION NUMBER: 09/370,265  
; PRIOR FILING DATE: 1999-08-09  
; PRIOR APPLICATION NUMBER: PCT/US99/16180  
; PRIOR FILING DATE: 1999-07-16  
; PRIOR APPLICATION NUMBER: 09/350836  
; PRIOR FILING DATE: 1999-07-09  
; PRIOR APPLICATION NUMBER: 09/273447  
; PRIOR FILING DATE: 1999-03-19  
; PRIOR APPLICATION NUMBER: 09/122449  
; PRIOR FILING DATE: 1998-07-24  
; PRIOR APPLICATION NUMBER: 09/244444  
; PRIOR FILING DATE: 1999-02-04  
; PRIOR APPLICATION NUMBER: 09/118,205  
; PRIOR FILING DATE: 1998-07-16  
; NUMBER OF SEQ ID NOS: 54  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 5  
; LENGTH: 428  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-286-926-5

Query Match 99.2%; Score 2235; DB 12; Length 428;  
Best Local Similarity 99.3%; Pred. No. 8.5e-217;

```

; PRIOR APPLICATION NUMBER: 60/263,800
; PRIOR FILING DATE: 2001-01-24
; PRIOR APPLICATION NUMBER: 60/269,942
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/286,183
; PRIOR FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: 60/313,627
; PRIOR FILING DATE: 2001-08-20
; PRIOR APPLICATION NUMBER: 60/318,712
; PRIOR FILING DATE: 2001-09-12
; NUMBER OF SEQ ID NOS: 292
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 127
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-231-913-127

Query Match          99.2%; Score 2235; DB 12; Length 428;
Best Local Similarity 99.3%; Pred. No. 8.5e-217;
Matches 425; Conservative 0; Mismatches 3; Indels 0; Gaps 0

Qy 1 MATSWGTVTFPMLVVSVCVSAVSHRNQQTWFEGLPSSMCPINVSASTLYGIMFDAGSTGT 60
Db 1 MATSWGTVTFPMLVVSVCVSAVSHRNQQTWFEGLPSSMCPINVSASTLYGIMFDAGSTGT 60

Qy 61 RIHYVTFVQKMPGLPILGEVEFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSI PRSHW 120
Db 61 RIHYVTFVQKMPGLPILGEVEFDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSI PRSHW 120

Qy 121 KKTPTVLLKATAGLRLLPEHKAKALLFEVKGIFRKSPLVPKGSVSIIMTQDEGIFAWVTY 180
Db 121 KKTPTVLLKATAGLRLLPEHKAKALLFEVKGIFRKSPLVPKGSVSIIMTQDEGIFAWVTY 180

```

```

161 NPLTQJGJGAKQZHVIVGLDGGKRSIQLTLPQFKNLEQJFPGJLDSFEMFNSIKUHIR 241
181 NFLTQJLGHGRQETVGTGLDGGASTQITFLPQPEKLTLEQTPRGYLTSFEMFNSYKLYTH 240
241 SYLGFGLKAARLATLGALETEGTGDGHTFRRSACLPRWLEAEWIFGVVKYQYGGNOGEVGF 300
241 SYLGFGLKAARLATLGALETEGTGDGHTFRRSACLPRWLEAEWIFGVVKYQYGGNOGEVGF 300
301 EPCVAEVLVRVVRGKLHQPEVQRGSFVAFSYYYDRAVDTMDIYEKGGLKVDFPERKAR 360
301 EPCVAEVLVRVVRGKLHQPEVQRGSFVAFSYYYDRAVDTMDIYEKGGLKVDFPERKAR 360
361 EVCNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTTKVNNIETGVALGNATHL 420
361 EVCNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTTKVNNIETGVALGNATHL 420
421 LOSLIGISH 428.
421 LOSLIGISH 428

```



;; PRIOR APPLICATION NUMBER: 09/122,449  
;; PRIOR FILING DATE: 1998-07-24  
;; PRIOR APPLICATION NUMBER: 09/244,444  
;; PRIOR FILING DATE: 1999-02-04  
;; NUMBER OF SEQ ID NOS: 23  
;; SOFTWARE: PatentIn Ver. 2.0  
;; SEQ ID NO 3  
;; LENGTH: 428  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-10-091-085-3

Query Match 99.2%; Score 2235; DB 14; Length 428;  
Best Local Similarity 99.3%; Pred. No. 8.5e-217;  
Matches 425; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 1 MATSGTVPFVFLVSVCSAVSHRNQQTWFEIGIFLSSMCPINVSASTLYGIMFDAGSTGT 60  
DB 1 MATSGTVPFVFLVSVCSAVSHRNQQTWFEIGIFLSSMCPINVSASTLYGIMFDAGSTGT 60  
QY 61 RIHVYTFVQKMPGQLPILGEVFDVSVKPGLSAFVDPKQGAETVQGLLEVAKDSIPRSHW 120  
DB 61 RIHVYTFVQKMPGQLPILGEVFDVSVKPGLSAFVDPKQGAETVQGLLEVAKDSIPRSHW 120  
QY 121 KKTVPVLKATAGLRLLEPHKAKALLFEVKEIFRKSPLVPGKSVSMTGDEGIFAWVTV 180  
DB 121 KKTVPVLKATAGLRLLEPHKAKALLFEVKEIFRKSPLVPGKSVSMTGDEGIFAWVTV 180  
QY 181 NFLTGQHGHRQETVGTDLGGASTQITFLPQPEKTLQTPRGYLTSPFEMFNSTYKLYTH 240  
DB 181 NFLTGQHGHRQETVGTDLGGASTQITFLPQPEKTLQTPRGYLTSPFEMFNSTYKLYTH 240  
QY 241 SYLGFGLKAARLATLGALETGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEVEVGF 300  
DB 241 SYLGFGLKAARLATLGALETGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEVEVGF 300  
QY 301 EPCYAEVLVRVGRKLGHPQEEVQSGSFYAFSYYYDRAVDTMDIDYKGGILKVEDFERKAR 360  
DB 301 EPCYAEVLVRVGRKLGHPQEEVQSGSFYAFSYYYDRAVDTMDIDYKGGILKVEDFERKAR 360  
QY 361 EVCNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQTLTKVNNIETGVALGATFHL 420  
DB 361 EVCNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQTLTKVNNIETGVALGATFHL 420  
QY 421 LQSLGISH 428  
DB 421 LQSLGISH 428

## RESULT 8

US-10-091-085-5  
;; Sequence 5, Application US/10091085  
;; Publication No. US2002014672A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Ford, John  
;; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE POLYPEPTIDES  
;; FILE REFERENCE: 28110/35761  
;; CURRENT APPLICATION NUMBER: US/10/091,085  
;; PRIOR FILING DATE: 2002-03-05  
;; PRIOR APPLICATION NUMBER: 09/350,836  
;; PRIOR FILING DATE: 1999-07-09  
;; PRIOR APPLICATION NUMBER: 09/273,447  
;; PRIOR FILING DATE: 1999-03-19  
;; PRIOR APPLICATION NUMBER: 09/118,205  
;; PRIOR FILING DATE: 1998-07-16  
;; PRIOR APPLICATION NUMBER: 09/122,449  
;; PRIOR FILING DATE: 1998-07-24  
;; PRIOR APPLICATION NUMBER: 09/244,444  
;; PRIOR FILING DATE: 1999-02-04  
;; NUMBER OF SEQ ID NOS: 23  
;; SOFTWARE: PatentIn Ver. 2.0

;; SEQ ID NO 5  
;; LENGTH: 428  
;; TYPE: PRT  
;; ORGANISM: Homo sapiens  
US-10-091-085-5

Query Match 99.2%; Score 2235; DB 14; Length 428;  
Best Local Similarity 99.3%; Pred. No. 8.5e-217;  
Matches 425; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 1 MATSGTVPFVFLVSVCSAVSHRNQQTWFEIGIFLSSMCPINVSASTLYGIMFDAGSTGT 60  
DB 1 MATSGTVPFVFLVSVCSAVSHRNQQTWFEIGIFLSSMCPINVSASTLYGIMFDAGSTGT 60  
QY 61 RIHVYTFVQKMPGQLPILGEVFDVSVKPGLSAFVDPKQGAETVQGLLEVAKDSIPRSHW 120  
DB 61 RIHVYTFVQKMPGQLPILGEVFDVSVKPGLSAFVDPKQGAETVQGLLEVAKDSIPRSHW 120  
QY 121 KKTVPVLKATAGLRLLEPHKAKALLFEVKEIFRKSPLVPGKSVSMTGDEGIFAWVTV 180  
DB 121 KKTVPVLKATAGLRLLEPHKAKALLFEVKEIFRKSPLVPGKSVSMTGDEGIFAWVTV 180  
QY 181 NFLTGQHGHRQETVGTDLGGASTQITFLPQPEKTLQTPRGYLTSPFEMFNSTYKLYTH 240  
DB 181 NFLTGQHGHRQETVGTDLGGASTQITFLPQPEKTLQTPRGYLTSPFEMFNSTYKLYTH 240  
QY 241 SYLGFGLKAARLATLGALETGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEVEVGF 300  
DB 241 SYLGFGLKAARLATLGALETGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEVEVGF 300  
QY 301 EPCYAEVLVRVGRKLGHPQEEVQSGSFYAFSYYYDRAVDTMDIDYKGGILKVEDFERKAR 360  
DB 301 EPCYAEVLVRVGRKLGHPQEEVQSGSFYAFSYYYDRAVDTMDIDYKGGILKVEDFERKAR 360  
QY 361 EVCNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQTLTKVNNIETGVALGATFHL 420  
DB 361 EVCNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQTLTKVNNIETGVALGATFHL 420  
QY 421 LQSLGISH 428  
DB 421 LQSLGISH 428

## RESULT 9

US-10-092-063-3  
;; Sequence 3, Application US/10092063  
;; Publication No. US20020173005A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Ford, John  
;; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE POLYPEPTIDES  
;; FILE REFERENCE: 28110/35908  
;; CURRENT APPLICATION NUMBER: US/10/092,063  
;; PRIOR FILING DATE: 2002-03-05  
;; PRIOR APPLICATION NUMBER: 09/370,265  
;; PRIOR FILING DATE: 2002-01-31  
;; PRIOR APPLICATION NUMBER: PCT/US99/16180  
;; PRIOR FILING DATE: 1999-07-16  
;; PRIOR APPLICATION NUMBER: 09/350,836  
;; PRIOR FILING DATE: 1999-07-09  
;; PRIOR APPLICATION NUMBER: 09/273,447  
;; PRIOR FILING DATE: 1999-03-19  
;; PRIOR APPLICATION NUMBER: 09/244,444  
;; PRIOR FILING DATE: 1999-02-04  
;; PRIOR APPLICATION NUMBER: 09/122,449  
;; PRIOR FILING DATE: 1998-07-24  
;; PRIOR APPLICATION NUMBER: 09/118,205  
;; PRIOR FILING DATE: 1998-07-16  
;; NUMBER OF SEQ ID NOS: 39  
;; SOFTWARE: PatentIn Ver. 2.0  
;; SEQ ID NO 3  
;; LENGTH: 428  
;; TYPE: PRT

ORGANISM: Homo sapiens  
US-10-092-063-3

Query Match 99.2%; Score 2235; DB 14; Length 428;  
Best Local Similarity 99.3%; Pred. No. 8.5e-217;  
Matches 425; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MATSGTGVFFMLVSVCSAVSHRNQQTWFEGLFLSSMCPINVSASTLYGIMFDAGSTGT 60  
DB 1 MATSGTGVFFMLVSVCSAVSHRNQQTWFEGLFLSSMCPINVSASTLYGIMFDAGSTGT 60

QY 61 RIHYVTFVQKMPGQLPILEGVFDVSKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120  
DB 61 RIHYVTFVQKMPGQLPILEGVFDVSKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120

QY 121 KKTVPVLKATAGLRLLEPHKAKALLFEVKEIFRKSPLVPKGSVIMTGDGEGIFAWTV 180  
DB 121 KKTVPVLKATAGLRLLEPHKAKALLFEVKEIFRKSPLVPKGSVIMTGDGEGIFAWTV 180

QY 181 NFLTGQLHGHROETVGTDLGGASTQITFLPQPEKTLQTPRGYLTSPFEMFNSTYKLYTH 240  
DB 181 NFLTGQLHGHROETVGTDLGGASTQITFLPQPEKTLQTPRGYLTSPFEMFNSTYKLYTH 240

QY 241 SYLGFGGLKAARLATLGALETGTGHTFRSACLPRWLEAEWIFGGVKYQYGGNOEGEVGF 300  
DB 241 SYLGFGGLKAARLATLGALETGTGHTFRSACLPRWLEAEWIFGGVKYQYGGNOEGEVGF 300

QY 301 EPCYAEVLVRVVRGKLHQPVEVQSGSFYAFSYYYDRAVDTMDIDYKGGILKVEDFERKAR 360  
DB 301 EPCYAEVLVRVVRGKLHQPVEVQSGSFYAFSYYYDRAVDTMDIDYKGGILKVEDFERKAR 360

QY 361 EVCNLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQTKKVNNIETGVALGATFHL 420  
DB 361 EVCNLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQTKKVNNIETGVALGATFHL 420

QY 421 LQSLGISH 428  
DB 421 LQSLGISH 428

RESULT 11  
US-10-286-926-25  
; Sequence 25, Application US/10286926  
; Publication No. US200301752A1  
; GENERAL INFORMATION:  
; APPLICANT: Ford, John  
; APPLICANT: Mulero, Julio  
; APPLICANT: Yeung, George  
; TITLE OF INVENTION: Methods and Materials Relating to CD39-Like  
; TITLE OF INVENTION: Polypeptides  
; FILE REFERENCE: 28110/36457CON  
; CURRENT APPLICATION NUMBER: US/10/286,926  
; CURRENT FILING DATE: 2002-11-01  
; PRIOR APPLICATION NUMBER: 09/557,800  
; PRIOR FILING DATE: 2000-04-25  
; PRIOR APPLICATION NUMBER: 09/481,238  
; PRIOR FILING DATE: 2000-01-11  
; PRIOR APPLICATION NUMBER: 09/370,265  
; PRIOR FILING DATE: 1999-08-09  
; PRIOR APPLICATION NUMBER: PCT/US99/16180  
; PRIOR FILING DATE: 1999-07-16  
; PRIOR APPLICATION NUMBER: 09/350836  
; PRIOR FILING DATE: 1999-07-09  
; PRIOR APPLICATION NUMBER: 09/273447  
; PRIOR FILING DATE: 1999-03-19  
; PRIOR APPLICATION NUMBER: 09/122449  
; PRIOR FILING DATE: 1998-07-24  
; PRIOR APPLICATION NUMBER: 09/244444  
; PRIOR FILING DATE: 1999-02-04  
; PRIOR APPLICATION NUMBER: 09/118,205  
; PRIOR FILING DATE: 1998-07-16  
; NUMBER OF SEQ ID NOS: 54  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 25  
; LENGTH: 405  
; TYPE: PRT

ORGANISM: Homo sapiens  
US-10-092-063-3

Query Match 99.2%; Score 2235; DB 14; Length 428;  
Best Local Similarity 99.3%; Pred. No. 8.5e-217;  
Matches 425; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MATSGTGVFFMLVSVCSAVSHRNQQTWFEGLFLSSMCPINVSASTLYGIMFDAGSTGT 60  
DB 1 MATSGTGVFFMLVSVCSAVSHRNQQTWFEGLFLSSMCPINVSASTLYGIMFDAGSTGT 60

QY 61 RIHYVTFVQKMPGQLPILEGVFDVSKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120  
DB 61 RIHYVTFVQKMPGQLPILEGVFDVSKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120

QY 121 KKTVPVLKATAGLRLLEPHKAKALLFEVKEIFRKSPLVPKGSVIMTGDGEGIFAWTV 180  
DB 121 KKTVPVLKATAGLRLLEPHKAKALLFEVKEIFRKSPLVPKGSVIMTGDGEGIFAWTV 180

QY 181 NFLTGQLHGHROETVGTDLGGASTQITFLPQPEKTLQTPRGYLTSPFEMFNSTYKLYTH 240  
DB 181 NFLTGQLHGHROETVGTDLGGASTQITFLPQPEKTLQTPRGYLTSPFEMFNSTYKLYTH 240

QY 241 SYLGFGGLKAARLATLGALETGTGHTFRSACLPRWLEAEWIFGGVKYQYGGNOEGEVGF 300  
DB 241 SYLGFGGLKAARLATLGALETGTGHTFRSACLPRWLEAEWIFGGVKYQYGGNOEGEVGF 300

QY 301 EPCYAEVLVRVVRGKLHQPVEVQSGSFYAFSYYYDRAVDTMDIDYKGGILKVEDFERKAR 360  
DB 301 EPCYAEVLVRVVRGKLHQPVEVQSGSFYAFSYYYDRAVDTMDIDYKGGILKVEDFERKAR 360

QY 361 EVCNLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQTKKVNNIETGVALGATFHL 420  
DB 361 EVCNLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQTKKVNNIETGVALGATFHL 420

QY 421 LQSLGISH 428  
DB 421 LQSLGISH 428

RESULT 10  
US-10-092-063-5  
; Sequence 5, Application US/10092063  
; Publication No. US20020173005A1  
; GENERAL INFORMATION:  
; APPLICANT: Ford, John  
; APPLICANT: Mulero, Julio  
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE POLYPEPTIDES  
; FILE REFERENCE: 28110/35908  
; CURRENT APPLICATION NUMBER: US/10/092,063  
; CURRENT FILING DATE: 2002-03-05  
; PRIOR APPLICATION NUMBER: 09/370,265  
; PRIOR FILING DATE: 2002-01-31  
; PRIOR APPLICATION NUMBER: PCT/US99/16180  
; PRIOR FILING DATE: 1999-07-16  
; PRIOR APPLICATION NUMBER: 09/350,836  
; PRIOR FILING DATE: 1999-07-09  
; PRIOR APPLICATION NUMBER: 09/273,447  
; PRIOR FILING DATE: 1999-03-19  
; PRIOR APPLICATION NUMBER: 09/244,444  
; PRIOR FILING DATE: 1999-02-04  
; PRIOR APPLICATION NUMBER: 09/122,449  
; PRIOR FILING DATE: 1998-07-24  
; PRIOR APPLICATION NUMBER: 09/118,205  
; PRIOR FILING DATE: 1998-07-16  
; NUMBER OF SEQ ID NOS: 39  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 5  
; LENGTH: 428  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-092-063-5

ORGANISM: Homo sapiens  
US-10-286-926-25

Query Match 92.8%; Score 2089; DB 12; Length 405;  
Best Local Similarity 99.2%; Pred. No. 4.4e-202;  
Matches 397; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MATSWGTVPFMLVSVCSAVSHRNQOTWFEGLFLSSMCPINVSASTLYGIMFDAGSTGT 60  
DB 1 MATSWGTVPFMLVSVCSAVSHRNQOTWFEGLFLSSMCPINVSASTLYGIMFDAGSTGT 60

QY 61 RIHVYTFVQKPGQLPILGEVDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120  
DB 61 RIHVYTFVQKPGQLPILGEVDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120

QY 121 KKTTPVVLKATAGRLLLPEHKAKALLFEVKEIFRKSPFLVPKGSVIMTQDEGIFAWTV 180  
DB 121 KKTTPVVLKATAGRLLLPEHKAKALLFEVKEIFRKSPFLVPKGSVIMTQDEGIFAWTV 180

QY 181 NFLTQGLHGRQETVGTDLGGASTQITFLPQFEKTLQTPRGYLTSPFEMFNSTYKLYTH 240  
DB 181 NFLTQGLHGRQETVGTDLGGASTQITFLPQFEKTLQTPRGYLTSPFEMFNSTYKLYTH 240

QY 241 SYLGFGKLAARLATLGALETGDTGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEVEGF 300  
DB 241 SYLGFGKLAARLATLGALETGDTGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEVEGF 300

QY 301 EPCYAEVLVRVGRKHLQHPVEVQSGFYAFSYYVDRAVDTMDIDYKGGILKVEDFERKAR 360  
DB 301 EPCYAEVLVRVGRKHLQHPVEVQSGFYAFSYYVDRAVDTMDIDYKGGILKVEDFERKAR 360

QY 361 EVCNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQ 400  
DB 361 EVCNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQ 400

RESULT 12  
US-10-092-063-25  
; Sequence 25, Application US/10092063  
; Publication No. US20020173005A1  
; GENERAL INFORMATION:  
; APPLICANT: Ford, John  
; APPLICANT: Mulero, Julio  
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE POLYPEPTIDES  
; FILE REFERENCE: 28110/35908  
; CURRENT APPLICATION NUMBER: US/10/092,063  
; CURRENT FILING DATE: 2002-03-05  
; PRIOR APPLICATION NUMBER: 09/370,265  
; PRIOR FILING DATE: 2002-01-31  
; PRIOR APPLICATION NUMBER: PCT/US99/16180  
; PRIOR FILING DATE: 1999-07-16  
; PRIOR APPLICATION NUMBER: 09/350,836  
; PRIOR FILING DATE: 1999-07-09  
; PRIOR APPLICATION NUMBER: 09/273,447  
; PRIOR FILING DATE: 1999-03-19  
; PRIOR APPLICATION NUMBER: 09/244,444  
; PRIOR FILING DATE: 1999-02-04  
; PRIOR APPLICATION NUMBER: 09/122,449  
; PRIOR FILING DATE: 1998-07-24  
; PRIOR APPLICATION NUMBER: 09/118,205  
; PRIOR FILING DATE: 1998-07-16  
; NUMBER OF SEQ ID NOS: 39  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 25  
; LENGTH: 405  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-092-063-25

Query Match 92.8%; Score 2089; DB 14; Length 405;  
Best Local Similarity 99.2%; Pred. No. 4.4e-202;  
Matches 397; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MATSWGTVPFMLVSVCSAVSHRNQOTWFEGLFLSSMCPINVSASTLYGIMFDAGSTGT 60  
DB 1 MATSWGTVPFMLVSVCSAVSHRNQOTWFEGLFLSSMCPINVSASTLYGIMFDAGSTGT 60

QY 61 RIHVYTFVQKPGQLPILGEVDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120  
DB 61 RIHVYTFVQKPGQLPILGEVDSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120

QY 121 KKTTPVVLKATAGRLLLPEHKAKALLFEVKEIFRKSPFLVPKGSVIMTQDEGIFAWTV 180  
DB 121 KKTTPVVLKATAGRLLLPEHKAKALLFEVKEIFRKSPFLVPKGSVIMTQDEGIFAWTV 180

QY 181 NFLTQGLHGRQETVGTDLGGASTQITFLPQFEKTLQTPRGYLTSPFEMFNSTYKLYTH 240  
DB 181 NFLTQGLHGRQETVGTDLGGASTQITFLPQFEKTLQTPRGYLTSPFEMFNSTYKLYTH 240

QY 241 SYLGFGKLAARLATLGALETGDTGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEVEGF 300  
DB 241 SYLGFGKLAARLATLGALETGDTGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEVEGF 300

QY 301 EPCYAEVLVRVGRKHLQHPVEVQSGFYAFSYYVDRAVDTMDIDYKGGILKVEDFERKAR 360  
DB 301 EPCYAEVLVRVGRKHLQHPVEVQSGFYAFSYYVDRAVDTMDIDYKGGILKVEDFERKAR 360

QY 361 EVCNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQ 400  
DB 361 EVCNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQ 400

RESULT 13  
US-10-231-913-126  
; Sequence 126, Application US/10231913  
; Publication No. US20040005576A1  
; GENERAL INFORMATION:  
; APPLICANT: Guo, Xiaojia S.  
; APPLICANT: Li, Li  
; APPLICANT: Patturajan, Meera  
; APPLICANT: Shimkets, Richard A.  
; APPLICANT: Casman, Stacie J.  
; APPLICANT: Malyanker, Uriel M.  
; APPLICANT: Tchernev, Velizar T.  
; APPLICANT: Vernet, Corine A.  
; APPLICANT: Spytek, Kimberly A.  
; APPLICANT: Shenoy, Suresh G.  
; APPLICANT: Alsobrook II, John P.  
; APPLICANT: Edinger, Schlomit  
; APPLICANT: Peyman, John A.  
; APPLICANT: Stone, David J.  
; APPLICANT: Ellerman, Karen  
; APPLICANT: Gangolli, Baha A.  
; APPLICANT: Boldog, Ference L.  
; APPLICANT: Colman, Steven D.  
; APPLICANT: Eisen, Andrew J.  
; APPLICANT: Liu, Xiaohong  
; APPLICANT: Padigaru, Muralidhara  
; APPLICANT: Spaderna, Steven K.  
; APPLICANT: Zerhusen, Bryan D.  
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same  
; FILE REFERENCE: 21402-216  
; CURRENT APPLICATION NUMBER: US/10/231,913  
; CURRENT FILING DATE: 2002-08-30  
; PRIOR APPLICATION NUMBER: 60/251,660  
; PRIOR FILING DATE: 2000-12-06  
; PRIOR APPLICATION NUMBER: 60/255,029  
; PRIOR FILING DATE: 2000-12-12  
; PRIOR APPLICATION NUMBER: 60/260,326  
; PRIOR FILING DATE: 2001-01-08  
; PRIOR APPLICATION NUMBER: 60/263,800  
; PRIOR FILING DATE: 2001-01-24  
; PRIOR APPLICATION NUMBER: 60/269,942  
; PRIOR FILING DATE: 2001-02-20  
; PRIOR APPLICATION NUMBER: 60/286,183  
; PRIOR FILING DATE: 2001-04-24

RESULT 14  
US-10-092-063-39  
; Sequence 39, Application US/10092063  
; Publication No. US20020173005A1  
; GENERAL INFORMATION:  
; APPLICANT: Ford, John  
; APPLICANT: Mulero, Julio  
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE POLYPEPTIDES  
; FILE REFERENCE: 28110/35908  
; CURRENT APPLICATION NUMBER: US/10/092,063  
; CURRENT FILING DATE: 2002-03-05  
; PRIOR APPLICATION NUMBER: 09/370,265  
; PRIOR FILING DATE: 2002-01-31  
; PRIOR APPLICATION NUMBER: PCT/US99/16180  
; PRIOR FILING DATE: 1999-07-16  
; PRIOR APPLICATION NUMBER: 09/350,836  
; PRIOR FILING DATE: 1999-07-09  
; PRIOR APPLICATION NUMBER: 09/273,447  
; PRIOR FILING DATE: 1999-03-19  
; PRIOR APPLICATION NUMBER: 09/244,444  
; PRIOR FILING DATE: 1999-02-04  
; PRIOR APPLICATION NUMBER: 09/122,449  
; PRIOR FILING DATE: 1998-07-24  
; PRIOR APPLICATION NUMBER: 09/118,205

```

RESULT 15
US-09-925-299-876
; Sequence 876, Application US/09925299
; Patent No. US20020055627A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA102
; CURRENT APPLICATION NUMBER: US/09/925,299
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05883
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1556
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 876
; LENGTH: 330
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (97)
; OTHER INFORMATION: Xaa equals any of the naturally occurring
; NAME/KEY: SITE
; LOCATION: (106)
; OTHER INFORMATION: Xaa equals any of the naturally occurring
; NAME/KEY: SITE
; LOCATION: (124)
; OTHER INFORMATION: Xaa equals any of the naturally occurring
; NAME/KEY: SITE

```

; LOCATION: (138)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (174)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (178)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
; NAME/KEY: SITE  
; LOCATION: (194)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-925-299-876

Query Match 73.0%; Score 1645; DB 9; Length 330;  
Best Local Similarity 96.3%; Pred. No. 2.3e-157;  
Matches 316; Conservative 0; Mismatches 12; Indels 0; Gaps 0;  
QY 101 AETVQGLLEVAKDSIPRSHWKTTPVVLKATAGLRLPEHKAKALLFEVKEIFRKSPFLVP 160  
Db 3 ARAVQGLLEVAKDSIPRSHWKTTPVVLKATAGLRLPEHKAKALLFEVKEIFRKSPFLVP 62  
QY 161 KGSVSIMTGODEGIFAWVTWNFLTQOLHGRQETVGTDLGGASTQITFLPOFEKLEQT 220  
Db 63 KGSVSIMDSDEGILAWVTWNFLTQOLHGRQETXGTDLGGAXTQITFLPOFEKLEQT 122  
QY 221 PRGYLTSFEMFNSTYKLYTHSYLGFLKAARLATLGALETEGTDGHTFRSACLPRWLEAE 280  
Db 123 PXGYLTSFEMFNSTYKLYTHSYLGFLKAARLATLGALETEGTDGHTFRSAXLPRXLEAE 182  
QY 281 WIFGGVKYQYGNQGEVGFPCYAEVLVRVGRKLHQPEVQVQSGSFYAFSYYDRAVDTD 340  
Db 183 WIFGGVKYQYGNQGEVGFPCYAEVLVRVGRKLHQPEVQVQSGSFYAFSYYDRAVDTD 242  
QY 341 MIDYKGGILKVEDFERKAREVCNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQ 400  
Db 243 MIDYKGGILKVEDFERKAREVCNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQ 302  
QY 401 LTKKVNNIETGWALGATFHLQSLGISH 428  
Db 303 LTKKVNNIETGWALGATFHLQSLGISH 330

Search completed: January 29, 2004, 10:14:50  
Job time : 30.5 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 10:05:35 ; Search time 29.5 Seconds  
(without alignments)  
3015.210 Million cell updates/sec

Title: US-10-091-085-3  
Perfect score: 2250  
Sequence: 1 MATSWGTVFMLVVCVCSA.....ETGALGATPHLLQSLGISH 428

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 789580 seqs, 207824079 residues

Total number of hits satisfying chosen parameters: 789580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2250	100.0	428	12	US-10-286-926-3
2	2250	100.0	428	12	US-10-286-926-5
3	2250	100.0	428	12	US-10-231-913-127
4	2250	100.0	428	14	US-10-091-085-3
5	2250	100.0	428	14	US-10-091-085-5
6	2250	100.0	428	14	US-10-092-063-3
7	2250	100.0	428	14	US-10-092-063-7
8	2235	99.3	428	12	US-10-286-926-7
9	2235	99.3	428	14	US-10-091-085-7
10	2235	99.3	428	14	US-10-092-063-7
11	2104	93.5	405	12	US-10-286-926-25
12	2104	93.5	405	14	US-10-092-063-25
13	1994.5	88.6	427	12	US-10-231-913-126
14	1837.5	81.7	465	14	US-10-092-063-39
15	1660	73.8	330	9	US-09-925-299-876

16	1660	73.8	330	11	US-09-925-299-876	Sequence 876, App
17	999	44.4	484	12	US-10-286-926-27	Sequence 27, Appl
18	999	44.4	484	12	US-10-231-913-123	Sequence 123, App
19	999	44.4	484	14	US-10-092-063-27	Sequence 27, Appl
20	996	44.3	467	12	US-10-231-913-36	Sequence 36, Appl
21	993.5	44.2	379	12	US-10-231-913-271	Sequence 271, App
22	992	44.1	484	12	US-10-231-913-124	Sequence 124, App
23	989	44.0	446	12	US-10-231-913-38	Sequence 38, Appl
24	986	43.8	455	12	US-10-231-913-125	Sequence 125, App
25	616.5	27.4	479	12	US-10-369-493-6447	Sequence 6447, App
26	590	26.2	476	9	US-09-835-147-4	Sequence 4, Appli
27	538.5	23.9	556	12	US-10-369-493-2169	Sequence 2169, Ap
28	533	23.7	476	9	US-09-835-147-3	Sequence 3, Appli
29	507.5	22.6	459	9	US-09-129-112-9	Sequence 9, Appli
30	503.5	22.4	467	9	US-09-129-112-19	Sequence 19, Appl
31	503	22.4	410	12	US-10-231-913-272	Sequence 272, App
32	498	22.1	462	9	US-09-129-112-2	Sequence 2, Appli
33	491.5	21.8	462	9	US-09-129-112-15	Sequence 15, Appl
34	483.5	21.5	518	12	US-10-369-493-1713	Sequence 1713, Ap
35	435	19.3	455	12	US-10-259-165-286	Sequence 286, App
36	429.5	19.1	459	12	US-10-259-165-74	Sequence 74, Appl
37	427	19.0	605	12	US-10-369-493-3333	Sequence 3333, Ap
38	387.5	17.2	485	12	US-10-369-493-6055	Sequence 6055, Ap
39	372.5	16.6	557	12	US-10-369-493-6789	Sequence 6789, Ap
40	372	16.5	630	12	US-10-369-493-1722	Sequence 1722, Ap
41	367	16.3	495	9	US-09-823-356-4	Sequence 4, Appli
42	358.5	15.9	458	12	US-10-199-672-496	Sequence 496, App
43	358.5	15.9	458	12	US-10-187-749-496	Sequence 496, App
44	358.5	15.9	458	12	US-10-194-457-496	Sequence 496, App
45	358.5	15.9	458	12	US-10-184-642-496	Sequence 496, App

## ALIGNMENTS

RESULT 1  
US-10-286-926-3  
; Sequence 3, Application US/10286926  
; Publication No. US20030175752A1  
; GENERAL INFORMATION:  
; APPLICANT: Ford, John  
; APPLICANT: Mulero, Julio  
; APPLICANT: Yeung, George  
; TITLE OF INVENTION: Methods and Materials Relating to CD39-Like  
; TITLE OF INVENTION: Polypeptides  
; FILE REFERENCE: 28110/36457CON  
; CURRENT APPLICATION NUMBER: US/10/286, 926  
; CURRENT FILING DATE: 2002-11-01  
; PRIOR APPLICATION NUMBER: 09/557,800  
; PRIOR FILING DATE: 2000-04-25  
; PRIOR APPLICATION NUMBER: 09/481,238  
; PRIOR FILING DATE: 2000-01-11  
; PRIOR APPLICATION NUMBER: 09/370,265  
; PRIOR FILING DATE: 1999-08-09  
; PRIOR APPLICATION NUMBER: PCT/US99/16180  
; PRIOR FILING DATE: 1999-07-16  
; PRIOR APPLICATION NUMBER: 09/350836  
; PRIOR FILING DATE: 1999-07-09  
; PRIOR APPLICATION NUMBER: 09/273447  
; PRIOR FILING DATE: 1999-03-19  
; PRIOR APPLICATION NUMBER: 09/122449  
; PRIOR FILING DATE: 1998-07-24  
; PRIOR APPLICATION NUMBER: 09/244444  
; PRIOR FILING DATE: 1999-02-04  
; PRIOR APPLICATION NUMBER: 09/118,205  
; PRIOR FILING DATE: 1998-07-16  
; NUMBER OF SEQ ID NOS: 54  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 428  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-286-926-3

Query Match 100.0%; Score 2250; DB 12; Length 428;  
Best Local Similarity 100.0%; Pred. No. 2.3e-215;  
Matches 428; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MATSGTGVFFMLVSVCSAVSHRNQQTWTFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60  
Db 1 MATSGTGVFFMLVSVCSAVSHRNQQTWTFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

Qy 61 RIHVTFVQKMPGQLPILEGEVFDVSKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120  
Db 61 RIHVTFVQKMPGQLPILEGEVFDVSKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120

Qy 121 KKTVPVLKATAGLRLLPEHKAKALLFEVKEIFRKSPLVPKGSVIMDGSDEGILAWTV 180  
Db 121 KKTVPVLKATAGLRLLPEHKAKALLFEVKEIFRKSPLVPKGSVIMDGSDEGILAWTV 180

Qy 181 NFLTGQLHGRQETVGTLDLGGASTQITFLPQPEKTLQTPRGYLTSEFNFNSTYKLYTH 240  
Db 181 NFLTGQLHGRQETVGTLDLGGASTQITFLPQPEKTLQTPRGYLTSEFNFNSTYKLYTH 240

Qy 241 SYLGFGKAAARLATLGALETGTDGHTFRSACLPRWLEAEWIFGVKYQYGGNQEGBVGF 300  
Db 241 SYLGFGKAAARLATLGALETGTDGHTFRSACLPRWLEAEWIFGVKYQYGGNQEGBVGF 300

Qy 301 EPCYAEVLVRVGRKHLQHPVEEVQSGSFYAFSYYYDRAVDTMDIDYKGGILKVEDFERKAR 360  
Db 301 EPCYAEVLVRVGRKHLQHPVEEVQSGSFYAFSYYYDRAVDTMDIDYKGGILKVEDFERKAR 360

Qy 361 EVCNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQTKKNNIETGWALGATFHL 420  
Db 361 EVCNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQTKKNNIETGWALGATFHL 420

Qy 421 LQSLGISH 428  
Db 421 LQSLGISH 428

RESULT 2  
US-10-286-926-5  
; Sequence 5, Application US/10286926  
; Publication No. US20030175752A1  
; GENERAL INFORMATION:  
; APPLICANT: Ford, John  
; APPLICANT: Mulero, Julio  
; APPLICANT: Yeung, George  
; TITLE OF INVENTION: Methods and Materials Relating to CD39-Like  
; FILE REFERENCE: 28110/36457CON  
; CURRENT FILING DATE: 2002-11-01  
; PRIOR FILING DATE: 2002-11-01  
; PRIOR FILING DATE: 2000-04-25  
; PRIOR FILING DATE: 2000-04-25  
; PRIOR FILING DATE: 2000-01-11  
; PRIOR FILING DATE: 1999-08-09  
; PRIOR FILING DATE: 1999-07-16  
; PRIOR FILING DATE: 1999-07-16  
; PRIOR FILING DATE: 1999-07-09  
; PRIOR FILING DATE: 1999-07-09  
; PRIOR FILING DATE: 1999-03-19  
; PRIOR FILING DATE: 1998-07-24  
; PRIOR FILING DATE: 1998-07-24  
; PRIOR FILING DATE: 1999-02-04  
; PRIOR FILING DATE: 1998-07-16  
; NUMBER OF SEQ ID NOS: 54  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 5  
; LENGTH: 428

; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-286-926-5

Query Match 100.0%; Score 2250; DB 12; Length 428;  
Best Local Similarity 100.0%; Pred. No. 2.3e-215;  
Matches 428; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MATSGTGVFFMLVSVCSAVSHRNQQTWTFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60  
Db 1 MATSGTGVFFMLVSVCSAVSHRNQQTWTFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

Qy 61 RIHVTFVQKMPGQLPILEGEVFDVSKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120  
Db 61 RIHVTFVQKMPGQLPILEGEVFDVSKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120

Qy 121 KKTVPVLKATAGLRLLPEHKAKALLFEVKEIFRKSPLVPKGSVIMDGSDEGILAWTV 180  
Db 121 KKTVPVLKATAGLRLLPEHKAKALLFEVKEIFRKSPLVPKGSVIMDGSDEGILAWTV 180

Qy 181 NFLTGQLHGRQETVGTLDLGGASTQITFLPQPEKTLQTPRGYLTSEFNFNSTYKLYTH 240  
Db 181 NFLTGQLHGRQETVGTLDLGGASTQITFLPQPEKTLQTPRGYLTSEFNFNSTYKLYTH 240

Qy 241 SYLGFGKAAARLATLGALETGTDGHTFRSACLPRWLEAEWIFGVKYQYGGNQEGBVGF 300  
Db 241 SYLGFGKAAARLATLGALETGTDGHTFRSACLPRWLEAEWIFGVKYQYGGNQEGBVGF 300

Qy 301 EPCYAEVLVRVGRKHLQHPVEEVQSGSFYAFSYYYDRAVDTMDIDYKGGILKVEDFERKAR 360  
Db 301 EPCYAEVLVRVGRKHLQHPVEEVQSGSFYAFSYYYDRAVDTMDIDYKGGILKVEDFERKAR 360

Qy 361 EVCNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQTKKNNIETGWALGATFHL 420  
Db 361 EVCNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQTKKNNIETGWALGATFHL 420

Qy 421 LQSLGISH 428  
Db 421 LQSLGISH 428

RESULT 3  
US-10-231-913-127  
; Sequence 127, Application US/10231913  
; Publication No. US20040005576A1  
; GENERAL INFORMATION:  
; APPLICANT: Guo, Xiaojia S.  
; APPLICANT: Li, Li  
; APPLICANT: Patturajan, Meera  
; APPLICANT: Shimkets, Richard A.  
; APPLICANT: Casman, Stacie J.  
; APPLICANT: Malyankar, Uriel M.  
; APPLICANT: Tchernev, Velizar T.  
; APPLICANT: Vernet, Corine A.  
; APPLICANT: Spytek, Kimberly A.  
; APPLICANT: Shenoy, Suresh G.  
; APPLICANT: Alsobrook II, John P.  
; APPLICANT: Bainger, Schlomit  
; APPLICANT: Peyman, John A.  
; APPLICANT: Stone, David J.  
; APPLICANT: Ellerman, Karen  
; APPLICANT: Gangolli, Esha A.  
; APPLICANT: Boldog, Ference L.  
; APPLICANT: Colman, Steven D.  
; APPLICANT: Eisen, Andrew J.  
; APPLICANT: Liu, Xiaohong  
; APPLICANT: Padigaru, Muralidhara  
; APPLICANT: Spaderna, Steven K.  
; APPLICANT: Zerhusen, Bryan D.  
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same  
; FILE REFERENCE: 21402-216  
; CURRENT APPLICATION NUMBER: US/10/231,913  
; CURRENT FILING DATE: 2002-08-30





; PRIOR APPLICATION NUMBER: 09/122,449  
 ; PRIOR FILING DATE: 1998-07-24  
 ; PRIOR APPLICATION NUMBER: 09/244,444  
 ; PRIOR FILING DATE: 1999-02-04  
 ; NUMBER OF SEQ ID NOS: 23  
 ; SOFTWARE: Patent In Ver. 2.0  
 ; SEQ ID NO 5  
 ; LENGTH: 428  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-10-091-085-5

Query Match 100.0%; Score 2250; DB 14; Length 428;  
 Best Local Similarity 100.0%; Pred. No. 2.3e-215;  
 Matches 428; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 MATSGTFFVFMVSVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60  
 Db 1 MATSGTFFVFMVSVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60  
 Qy 61 RIHYTVFVQKMPGQLPILEGVFDVSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120  
 Db 61 RIHYTVFVQKMPGQLPILEGVFDVSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120  
 Qy 121 KKTVPVLKATAGLRLPEHKAALLFEVKEIFRKSPLVPKGSVIMDGSDEGILAWTV 180  
 Db 121 KKTVPVLKATAGLRLPEHKAALLFEVKEIFRKSPLVPKGSVIMDGSDEGILAWTV 180  
 Qy 181 NFLTGQLHGHROETVGTDLGGASTQITFLPQFEKTLQTPRGYLTSEFNFNSTYKLYTH 240  
 Db 181 NFLTGQLHGHROETVGTDLGGASTQITFLPQFEKTLQTPRGYLTSEFNFNSTYKLYTH 240  
 Qy 241 SYLGFGGLKAARLATLGALETGTGHTFRSACLPRWLEAEWIFGGVYQYGGNOEGEVGF 300  
 Db 241 SYLGFGGLKAARLATLGALETGTGHTFRSACLPRWLEAEWIFGGVYQYGGNOEGEVGF 300  
 Qy 301 EPCVAELRVVRGKHLQHPVEVQSGSFYAFSYVDRAVDTMDIDYKGGILKVEDFERKAR 360  
 Db 301 EPCVAELRVVRGKHLQHPVEVQSGSFYAFSYVDRAVDTMDIDYKGGILKVEDFERKAR 360  
 Qy 361 EVCNLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQTKKVNNIETGWALGATFHL 420  
 Db 361 EVCNLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQTKKVNNIETGWALGATFHL 420  
 Qy 421 LQSLGISH 428  
 Db 421 LQSLGISH 428

RESULT 6  
 US-10-092-063-3  
 ; Sequence 3, Application US/10092063  
 ; Publication No. US20020173005A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ford, John  
 ; APPLICANT: Mulero, Julio  
 ; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE POLYPEPTIDES  
 ; FILE REFERENCE: 28110/35908  
 ; CURRENT APPLICATION NUMBER: US/10/092,063  
 ; CURRENT FILING DATE: 2002-03-05  
 ; PRIOR APPLICATION NUMBER: 09/370,265  
 ; PRIOR FILING DATE: 2002-01-31  
 ; PRIOR APPLICATION NUMBER: PCT/US99/16180  
 ; PRIOR FILING DATE: 1999-07-16  
 ; PRIOR APPLICATION NUMBER: 09/350,836  
 ; PRIOR FILING DATE: 1999-07-09  
 ; PRIOR APPLICATION NUMBER: 09/273,447  
 ; PRIOR FILING DATE: 1999-03-19  
 ; PRIOR APPLICATION NUMBER: 09/244,444  
 ; PRIOR FILING DATE: 1999-02-04  
 ; PRIOR APPLICATION NUMBER: 09/122,449  
 ; PRIOR FILING DATE: 1998-07-24  
 ; PRIOR APPLICATION NUMBER: 09/118,205

; PRIOR FILING DATE: 1998-07-16  
 ; NUMBER OF SEQ ID NOS: 39  
 ; SOFTWARE: Patent In Ver. 2.0  
 ; SEQ ID NO 3  
 ; LENGTH: 428  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-10-092-063-3

Query Match 100.0%; Score 2250; DB 14; Length 428;  
 Best Local Similarity 100.0%; Pred. No. 2.3e-215;  
 Matches 428; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 MATSGTFFVFMVSVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60  
 Db 1 MATSGTFFVFMVSVCSAVSHRNQQTWFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60  
 Qy 61 RIHYTVFVQKMPGQLPILEGVFDVSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120  
 Db 61 RIHYTVFVQKMPGQLPILEGVFDVSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120  
 Qy 121 KKTVPVLKATAGLRLPEHKAALLFEVKEIFRKSPLVPKGSVIMDGSDEGILAWTV 180  
 Db 121 KKTVPVLKATAGLRLPEHKAALLFEVKEIFRKSPLVPKGSVIMDGSDEGILAWTV 180  
 Qy 181 NFLTGQLHGHROETVGTDLGGASTQITFLPQFEKTLQTPRGYLTSEFNFNSTYKLYTH 240  
 Db 181 NFLTGQLHGHROETVGTDLGGASTQITFLPQFEKTLQTPRGYLTSEFNFNSTYKLYTH 240  
 Qy 241 SYLGFGGLKAARLATLGALETGTGHTFRSACLPRWLEAEWIFGGVYQYGGNOEGEVGF 300  
 Db 241 SYLGFGGLKAARLATLGALETGTGHTFRSACLPRWLEAEWIFGGVYQYGGNOEGEVGF 300  
 Qy 301 EPCVAELRVVRGKHLQHPVEVQSGSFYAFSYVDRAVDTMDIDYKGGILKVEDFERKAR 360  
 Db 301 EPCVAELRVVRGKHLQHPVEVQSGSFYAFSYVDRAVDTMDIDYKGGILKVEDFERKAR 360  
 Qy 361 EVCNLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQTKKVNNIETGWALGATFHL 420  
 Db 361 EVCNLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQTKKVNNIETGWALGATFHL 420  
 Qy 421 LQSLGISH 428  
 Db 421 LQSLGISH 428

RESULT 7  
 US-10-092-063-5  
 ; Sequence 5, Application US/10092063  
 ; Publication No. US20020173005A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ford, John  
 ; APPLICANT: Mulero, Julio  
 ; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE POLYPEPTIDES  
 ; FILE REFERENCE: 28110/35908  
 ; CURRENT APPLICATION NUMBER: US/10/092,063  
 ; CURRENT FILING DATE: 2002-03-05  
 ; PRIOR APPLICATION NUMBER: 09/370,265  
 ; PRIOR FILING DATE: 2002-01-31  
 ; PRIOR APPLICATION NUMBER: PCT/US99/16180  
 ; PRIOR FILING DATE: 1999-07-16  
 ; PRIOR APPLICATION NUMBER: 09/350,836  
 ; PRIOR FILING DATE: 1999-07-09  
 ; PRIOR APPLICATION NUMBER: 09/273,447  
 ; PRIOR FILING DATE: 1999-03-19  
 ; PRIOR APPLICATION NUMBER: 09/244,444  
 ; PRIOR FILING DATE: 1999-02-04  
 ; PRIOR APPLICATION NUMBER: 09/122,449  
 ; PRIOR FILING DATE: 1998-07-24  
 ; PRIOR APPLICATION NUMBER: 09/118,205  
 ; PRIOR FILING DATE: 1998-07-16  
 ; NUMBER OF SEQ ID NOS: 39  
 ; SOFTWARE: Patent In Ver. 2.0

```
; SEQ ID NO 5
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-092-063-5

Query Match      100.0%; Score 2250; DB 14; Length 428;
Best Local Similarity 100.0%; Pred. No. 2.3e-215;
Matches 428; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATSGTGVFFMLVSVCSAVSHRNQQTWFEIGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
DB 1 MATSGTGVFFMLVSVCSAVSHRNQQTWFEIGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
QY 61 RIHVYTFVQKMPGQLPILGEVFDVSKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
DB 61 RIHVYTFVQKMPGQLPILGEVFDVSKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
QY 121 KKTVPVVKATAGLRLLPEHKAKALLFEVKEIFPKSPFLVPKGSVSIIMDGSDEGILAWTV 180
DB 121 KKTVPVVKATAGLRLLPEHKAKALLFEVKEIFPKSPFLVPKGSVSIIMDGSDEGILAWTV 180
QY 181 NFLTQGLHGHROETVGTLDLGGASTQITFLPOFEKTLTQTPRGYLTSEFMFNSTYKLYTH 240
DB 181 NFLTQGLHGHROETVGTLDLGGASTQITFLPOFEKTLTQTPRGYLTSEFMFNSTYKLYTH 240
QY 241 SYLGFGKLAARLALGALTEGTGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEVEVGF 300
DB 241 SYLGFGKLAARLALGALTEGTGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEVEVGF 300
QY 301 EPCYAEVLVRGKLHQPVEVQSGSYAFSYYYDRAVDTMDIDYKGGILKVEDFERKAR 360
DB 301 EPCYAEVLVRGKLHQPVEVQSGSYAFSYYYDRAVDTMDIDYKGGILKVEDFERKAR 360
QY 361 EVCNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTCKVNNIETGHALGATPHL 420
DB 361 EVCNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTCKVNNIETGHALGATPHL 420
QY 421 LQSLGISH 428
DB 421 LQSLGISH 428

RESULT 8
US-10-286-926-7
; Sequence 7, Application US/10286926
; Publication No. US2003017572A1
; GENERAL INFORMATION:
; APPLICANT: Mulero, Julio
; APPLICANT: Yeung, George
; TITLE OF INVENTION: Methods and Materials Relating to CD39-Like
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 28110/36457CON
; CURRENT APPLICATION NUMBER: US/10/286,926
; CURRENT FILING DATE: 2002-11-01
; PRIOR APPLICATION NUMBER: 09/557,800
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/481,238
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: 09/370,265
; PRIOR FILING DATE: 1999-08-09
; PRIOR APPLICATION NUMBER: PCT/US99/16180
; PRIOR FILING DATE: 1999-07-16
; PRIOR APPLICATION NUMBER: 09/350836
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/273447
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: 09/122449
; PRIOR FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: 09/244444
; PRIOR FILING DATE: 1999-02-04
; PRIOR APPLICATION NUMBER: 09/118,205

; SEQ ID NO 7
; LENGTH: 428
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-286-926-7

Query Match      99.3%; Score 2235; DB 12; Length 428;
Best Local Similarity 99.3%; Pred. No. 7.3e-214;
Matches 425; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MATSGTGVFFMLVSVCSAVSHRNQQTWFEIGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
DB 1 MATSGTGVFFMLVSVCSAVSHRNQQTWFEIGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
QY 61 RIHVYTFVQKMPGQLPILGEVFDVSKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
DB 61 RIHVYTFVQKMPGQLPILGEVFDVSKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
QY 121 KKTVPVVKATAGLRLLPEHKAKALLFEVKEIFPKSPFLVPKGSVSIIMDGSDEGILAWTV 180
DB 121 KKTVPVVKATAGLRLLPEHKAKALLFEVKEIFPKSPFLVPKGSVSIIMDGSDEGILAWTV 180
QY 181 NFLTQGLHGHROETVGTLDLGGASTQITFLPOFEKTLTQTPRGYLTSEFMFNSTYKLYTH 240
DB 181 NFLTQGLHGHROETVGTLDLGGASTQITFLPOFEKTLTQTPRGYLTSEFMFNSTYKLYTH 240
QY 241 SYLGFGKLAARLALGALTEGTGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEVEVGF 300
DB 241 SYLGFGKLAARLALGALTEGTGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEVEVGF 300
QY 301 EPCYAEVLVRGKLHQPVEVQSGSYAFSYYYDRAVDTMDIDYKGGILKVEDFERKAR 360
DB 301 EPCYAEVLVRGKLHQPVEVQSGSYAFSYYYDRAVDTMDIDYKGGILKVEDFERKAR 360
QY 361 EVCNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTCKVNNIETGHALGATPHL 420
DB 361 EVCNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQLTCKVNNIETGHALGATPHL 420
QY 421 LQSLGISH 428
DB 421 LQSLGISH 428

RESULT 9
US-10-091-085-7
; Sequence 7, Application US/10091085
; Publication No. US2002014672A1
; GENERAL INFORMATION:
; APPLICANT: Ford, John
; APPLICANT: Mulero, Julio
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE
; TITLE OF INVENTION: POLYPEPTIDES
; FILE REFERENCE: 28110/35761
; CURRENT APPLICATION NUMBER: US/10/091,085
; CURRENT FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: 09/350,836
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/273,447
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: 09/118,205
; PRIOR FILING DATE: 1998-07-16
; PRIOR APPLICATION NUMBER: 09/122,449
; PRIOR FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: 09/244,444
; PRIOR FILING DATE: 1999-02-04
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 428
; TYPE: PRT
```

Query Match 99.3%; Score 2235; DB 14; Length 428;  
Best Local Similarity 99.3%; Pred. No. 7.3e-214;  
Matches 425; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MATSWGTFFMLVVCVCSAVSHRNQOTWPEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60  
DB 1 MATSWGTFFMLVVCVCSAVSHRNQOTWPEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

QY 61 RIHYVTFVQKMPGQLPILGEVFDVSKPGLSAFVDOPKQGAETVQGLLEVAKDSIPRSHW 120  
DB 61 RIHYVTFVQKMPGQLPILGEVFDVSKPGLSAFVDOPKQGAETVQGLLEVAKDSIPRSHW 120

QY 121 KKTVPVLKATAGLRLLEPHKAKALLFEVKEIFRKSPPFLVPKGSVIMDGSDEGILAWTV 180  
DB 121 KKTVPVLKATAGLRLLEPHKAKALLFEVKEIFRKSPPFLVPKGSVIMDGSDEGILAWTV 180

QY 181 NFLTGQLHGRHQETVGTDLGGASTQITFLPQPEKTLQTPRGYLTSEFENFNSTYKLYTH 240  
DB 181 NFLTGQLHGRHQETVGTDLGGASTQITFLPQPEKTLQTPRGYLTSEFENFNSTYKLYTH 240

QY 241 SYLGFGLKAARLATLGALETGTGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEVEVGF 300  
DB 241 SYLGFGLKAARLATLGALETGTGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEVEVGF 300

QY 301 EPCYAEVLVRVVRGKLHQPVEVQSGSFYAFSYYVDRAVDTMDIDYKGGILKVEDFERKAR 360  
DB 301 EPCYAEVLVRVVRGKLHQPVEVQSGSFYAFSYYVDRAVDTMDIDYKGGILKVEDFERKAR 360

QY 361 EVCNLENFTSGSPFLCNDLSYITALLKDGFGFADSTVLQITKKVNNIETGVALGATFHL 420  
DB 361 EVCNLENFTSGSPFLCNDLSYITALLKDGFGFADSTVLQITKKVNNIETGVALGATFHL 420

QY 421 LQSLGISH 428  
DB 421 LQSLGISH 428

RESULT 11  
US-10-286-926-25  
Sequence 25, Application US/10286926  
Publication No. US20030175752A1  
GENERAL INFORMATION:  
APPLICANT: Mulero, Julio  
APPLICANT: Ford, John  
TITLE OF INVENTION: Methods and Materials Relating to CD39-Like  
TITLE OF INVENTION: Polypeptides  
FILE REFERENCE: 28110/36457CON  
CURRENT APPLICATION NUMBER: US/10/286,926  
CURRENT FILING DATE: 2002-11-01  
PRIOR APPLICATION NUMBER: 09/557,800  
PRIOR FILING DATE: 2000-04-25  
PRIOR APPLICATION NUMBER: 09/481,238  
PRIOR FILING DATE: 2000-01-11  
PRIOR APPLICATION NUMBER: 09/370,265  
PRIOR FILING DATE: 1999-08-09  
PRIOR APPLICATION NUMBER: PCT/US99/16180  
PRIOR FILING DATE: 1999-07-16  
PRIOR APPLICATION NUMBER: 09/350836  
PRIOR FILING DATE: 1999-07-09  
PRIOR APPLICATION NUMBER: 09/273447  
PRIOR FILING DATE: 1999-03-19  
PRIOR APPLICATION NUMBER: 09/122449  
PRIOR FILING DATE: 1998-07-24  
PRIOR APPLICATION NUMBER: 09/244444  
PRIOR FILING DATE: 1999-02-04  
PRIOR APPLICATION NUMBER: 09/118,205  
PRIOR FILING DATE: 1998-07-16  
NUMBER OF SEQ ID NOS: 54  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 25  
LENGTH: 405  
TYPE: PRT

Query Match 99.3%; Score 2235; DB 14; Length 428;  
Best Local Similarity 99.3%; Pred. No. 7.3e-214;  
Matches 425; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 MATSWGTFFMLVVCVCSAVSHRNQOTWPEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60  
DB 1 MATSWGTFFMLVVCVCSAVSHRNQOTWPEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

QY 61 RIHYVTFVQKMPGQLPILGEVFDVSKPGLSAFVDOPKQGAETVQGLLEVAKDSIPRSHW 120  
DB 61 RIHYVTFVQKMPGQLPILGEVFDVSKPGLSAFVDOPKQGAETVQGLLEVAKDSIPRSHW 120

QY 121 KKTVPVLKATAGLRLLEPHKAKALLFEVKEIFRKSPPFLVPKGSVIMDGSDEGILAWTV 180  
DB 121 KKTVPVLKATAGLRLLEPHKAKALLFEVKEIFRKSPPFLVPKGSVIMDGSDEGILAWTV 180

QY 181 NFLTGQLHGRHQETVGTDLGGASTQITFLPQPEKTLQTPRGYLTSEFENFNSTYKLYTH 240  
DB 181 NFLTGQLHGRHQETVGTDLGGASTQITFLPQPEKTLQTPRGYLTSEFENFNSTYKLYTH 240

QY 241 SYLGFGLKAARLATLGALETGTGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEVEVGF 300  
DB 241 SYLGFGLKAARLATLGALETGTGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEVEVGF 300

QY 301 EPCYAEVLVRVVRGKLHQPVEVQSGSFYAFSYYVDRAVDTMDIDYKGGILKVEDFERKAR 360  
DB 301 EPCYAEVLVRVVRGKLHQPVEVQSGSFYAFSYYVDRAVDTMDIDYKGGILKVEDFERKAR 360

QY 361 EVCNLENFTSGSPFLCNDLSYITALLKDGFGFADSTVLQITKKVNNIETGVALGATFHL 420  
DB 361 EVCNLENFTSGSPFLCNDLSYITALLKDGFGFADSTVLQITKKVNNIETGVALGATFHL 420

QY 421 LQSLGISH 428  
DB 421 LQSLGISH 428

RESULT 10  
US-10-092-063-7  
Sequence 7, Application US/10092063  
Publication No. US20020173005A1  
GENERAL INFORMATION:  
APPLICANT: Ford, John  
APPLICANT: Mulero, Julio  
TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE POLYPEPTIDES  
FILE REFERENCE: 28110/35908  
CURRENT APPLICATION NUMBER: US/10/092,063  
CURRENT FILING DATE: 2002-03-05  
PRIOR APPLICATION NUMBER: 09/370,265  
PRIOR FILING DATE: 2002-01-31  
PRIOR APPLICATION NUMBER: PCT/US99/16180  
PRIOR FILING DATE: 1999-07-16  
PRIOR APPLICATION NUMBER: 09/350,836  
PRIOR FILING DATE: 1999-07-09  
PRIOR APPLICATION NUMBER: 09/273,447  
PRIOR FILING DATE: 1999-03-19  
PRIOR APPLICATION NUMBER: 09/244,444  
PRIOR FILING DATE: 1999-02-04  
PRIOR APPLICATION NUMBER: 09/122,449  
PRIOR FILING DATE: 1998-07-24  
PRIOR APPLICATION NUMBER: 09/118,205  
PRIOR FILING DATE: 1998-07-16  
NUMBER OF SEQ ID NOS: 39  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 7  
LENGTH: 428  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-092-063-7

ORGANISM: Homo sapiens  
US-10-286-926-25  
Query Match 93.5%; Score 2104; DB 12; Length 405;  
Best Local Similarity 100.0%; Pred. No. 7.7e-201;  
Matches 400; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATSWGTVFPMVLVSCVCSAVSHRNQOTWPEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60  
DB 1 MATSWGTVFPMVLVSCVCSAVSHRNQOTWPEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

QY 61 RIHVYTFVQKPGOLPILGEVFDVSKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120  
DB 61 RIHVYTFVQKPGOLPILGEVFDVSKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120

QY 121 KKTTPVVLKATAGRLRLPEHKAALLFEVKEIFRKSPLVPKGSVIMDGSDEGILAWTV 180  
DB 121 KKTTPVVLKATAGRLRLPEHKAALLFEVKEIFRKSPLVPKGSVIMDGSDEGILAWTV 180

QY 181 NFLTQGLHGRQETVGTDLGGASTQITFLPQPEKTLQTPRGYLTSPFNFNSTYKLYTH 240  
DB 181 NFLTQGLHGRQETVGTDLGGASTQITFLPQPEKTLQTPRGYLTSPFNFNSTYKLYTH 240

QY 241 SYLGFGKKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEVEVGF 300  
DB 241 SYLGFGKKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEVEVGF 300

QY 301 EPCYAEVLVRVGRKLGHPQEEVQSGSFYAFSYYYDRAVDTMDIDYKGGILKVEDFERKAR 360  
DB 301 EPCYAEVLVRVGRKLGHPQEEVQSGSFYAFSYYYDRAVDTMDIDYKGGILKVEDFERKAR 360

QY 361 EVCNLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQ 400  
DB 361 EVCNLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQ 400

RESULT 12  
US-10-092-063-25  
; Sequence 25, Application US/10092063  
; Publication No. US20020173005A1  
; GENERAL INFORMATION:  
; APPLICANT: Ford, John  
; APPLICANT: Mulero, Julio  
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE POLYPEPTIDES  
; FILE REFERENCE: 28110/35908  
; CURRENT APPLICATION NUMBER: US/10/092,063  
; CURRENT FILING DATE: 2002-03-05  
; PRIOR APPLICATION NUMBER: 09/370,265  
; PRIOR FILING DATE: 2002-01-31  
; PRIOR APPLICATION NUMBER: PCT/US99/16180  
; PRIOR FILING DATE: 1999-07-16  
; PRIOR APPLICATION NUMBER: 09/350,836  
; PRIOR FILING DATE: 1999-07-09  
; PRIOR APPLICATION NUMBER: 09/273,447  
; PRIOR FILING DATE: 1999-03-19  
; PRIOR APPLICATION NUMBER: 09/244,444  
; PRIOR FILING DATE: 1999-02-04  
; PRIOR APPLICATION NUMBER: 09/122,449  
; PRIOR FILING DATE: 1998-07-24  
; PRIOR APPLICATION NUMBER: 09/118,205  
; PRIOR FILING DATE: 1998-07-16  
; NUMBER OF SEQ ID NOS: 39  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 25  
; LENGTH: 405  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-092-063-25

Query Match 93.5%; Score 2104; DB 14; Length 405;  
Best Local Similarity 100.0%; Pred. No. 7.7e-201;  
Matches 400; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MATSWGTVFPMVLVSCVCSAVSHRNQOTWPEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60  
DB 1 MATSWGTVFPMVLVSCVCSAVSHRNQOTWPEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

QY 61 RIHVYTFVQKPGOLPILGEVFDVSKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120  
DB 61 RIHVYTFVQKPGOLPILGEVFDVSKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120

QY 121 KKTTPVVLKATAGRLRLPEHKAALLFEVKEIFRKSPLVPKGSVIMDGSDEGILAWTV 180  
DB 121 KKTTPVVLKATAGRLRLPEHKAALLFEVKEIFRKSPLVPKGSVIMDGSDEGILAWTV 180

QY 181 NFLTQGLHGRQETVGTDLGGASTQITFLPQPEKTLQTPRGYLTSPFNFNSTYKLYTH 240  
DB 181 NFLTQGLHGRQETVGTDLGGASTQITFLPQPEKTLQTPRGYLTSPFNFNSTYKLYTH 240

QY 241 SYLGFGKKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEVEVGF 300  
DB 241 SYLGFGKKAARLATLGALETEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGNQEVEVGF 300

QY 301 EPCYAEVLVRVGRKLGHPQEEVQSGSFYAFSYYYDRAVDTMDIDYKGGILKVEDFERKAR 360  
DB 301 EPCYAEVLVRVGRKLGHPQEEVQSGSFYAFSYYYDRAVDTMDIDYKGGILKVEDFERKAR 360

QY 361 EVCNLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQ 400  
DB 361 EVCNLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQ 400

RESULT 13  
US-10-231-913-126  
; Sequence 126, Application US/10231913  
; Publication No. US20040005576A1  
; GENERAL INFORMATION:  
; APPLICANT: Guo, Xiaojia S.  
; APPLICANT: Li, Li  
; APPLICANT: Patturajan, Meera  
; APPLICANT: Shimkete, Richard A.  
; APPLICANT: Casman, Stacie J.  
; APPLICANT: Malyankar, Uriel M.  
; APPLICANT: Tchernev, Velizar T.  
; APPLICANT: Vernet, Corine A.  
; APPLICANT: Spytek, Kimberly A.  
; APPLICANT: Shenoy, Suresh G.  
; APPLICANT: Alsobrook II, John P.  
; APPLICANT: Edinger, Schlomit  
; APPLICANT: Peyman, John A.  
; APPLICANT: Stone, David J.  
; APPLICANT: Ellerman, Karen  
; APPLICANT: Gangolli, Esha A.  
; APPLICANT: Boldog, Ference L.  
; APPLICANT: Colman, Steven D.  
; APPLICANT: Eissen, Andrew J.  
; APPLICANT: Liu, Xiaohong  
; APPLICANT: Padigaru, Muralidhara  
; APPLICANT: Spaderna, Steven K.  
; APPLICANT: Zerhusen, Bryan D.  
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same  
; FILE REFERENCE: 21402-216  
; CURRENT APPLICATION NUMBER: US/10/231,913  
; CURRENT FILING DATE: 2002-08-30  
; PRIOR APPLICATION NUMBER: 60/251,660  
; PRIOR FILING DATE: 2000-12-06  
; PRIOR APPLICATION NUMBER: 60/255,029  
; PRIOR FILING DATE: 2000-12-12  
; PRIOR APPLICATION NUMBER: 60/260,326  
; PRIOR FILING DATE: 2001-01-08  
; PRIOR APPLICATION NUMBER: 60/263,800  
; PRIOR FILING DATE: 2001-01-24  
; PRIOR APPLICATION NUMBER: 60/269,942  
; PRIOR FILING DATE: 2001-02-20  
; PRIOR APPLICATION NUMBER: 60/286,183  
; PRIOR FILING DATE: 2001-04-24

; PRIOR APPLICATION NUMBER: 60/313,627
; PRIOR FILING DATE: 2001-08-20
; PRIOR APPLICATION NUMBER: 60/318,712
; PRIOR FILING DATE: 2001-09-12
; NUMBER OF SEQ ID NOS: 292
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 126
; LENGTH: 427
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-231-913-126

Query Match 88.6%; Score 1994.5; DB 12; Length 427;
Best Local Similarity 88.3%; Pred. No. 6.9e-190;
Matches 377; Conservative 24; Mismatches 25; Indels 1; Gaps 1;

Qy 1 MATSGTGFVFMVSVCSAVSHRNQQTWTFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
Db 1 MATSGAV-FMLIIACVSGSTVFYREQQTWTFEGVFLSSMCPINVSAGTFYGMFDAGSTGT 59
61 RIHYVTFVQKMPGQLPILLEGVFSVSKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
Db 60 RIHYVTFVQKTAGQLPFLGFIPIFSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 119
121 KKTVPVLKATAGRLLPKHAALLPEVKEIFRSPFLVPKGSVIMDGSDEGILAWTV 180
Db 120 ERTPVWLKATAGRLLPKHAALLPEVKEIFRSPFLVPKGSVIMDGSDEGILAWTV 179
181 NFLTQGLHGRQETVGTDLGGASTQTITFLPQPEKTLTQTPRGYLTSEFMFNSYKLYTH 240
Db 180 NFLTQGLHGRQETVGTDLGGASTQTITFLPQPEKTLTQTPRGYLTSEFMFNSYKLYTH 239
241 SYLGFGKLAARLATLGALETEGTDGHTFRSACLPRMLEAEWIFGVVYQYGGNOEGEVSF 300
Db 240 SYLGFGKLAARLATLGALETEGTDGHTFRSACLPRMLEAEWIFGVVYQYGGNOEGEVSF 299
301 EPCYAEVLVRVVRGKLHQPEEVQSGSFYAFSYYVDRAVDTMDIDYKGGILKVEDFERKAR 360
Db 300 EPCYAEVLVRVVRGKLHQPEEVQSGSFYAFSYYVDRAVDTMDIDYKGGILKVEDFERKAR 359
361 EVCNLENFTSGSPFLCNDLSYITALLKDGFGFADSTVLQITKKVNNIETGVALGATFHL 420
Db 360 EVCNLDGSGSGSPFLCNDLTYITALLKDGFGFADSTVLQITKKVNNIETGVALGATFHL 419
421 LQSLGIS 427
Db 420 LQSLGIT 426

RESULT 14
US-10-092-063-39
; Sequence 39, Application US/10092063
; Patent No. US20020173005A1
; GENERAL INFORMATION:
; APPLICANT: Mulero, Julio
; APPLICANT: Ford, John
; TITLE OF INVENTION: METHODS AND MATERIALS RELATING TO NOVEL CD39-LIKE POLYPEPTIDES
; FILE REFERENCE: 28110/35908
; CURRENT APPLICATION NUMBER: US/10/092,063
; PRIOR FILING DATE: 2002-03-05
; PRIOR APPLICATION NUMBER: 09/370,265
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: PCT/US99/16180
; PRIOR FILING DATE: 1999-07-16
; PRIOR APPLICATION NUMBER: 09/350,836
; PRIOR FILING DATE: 1999-07-09
; PRIOR APPLICATION NUMBER: 09/273,447
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: 09/244,444
; PRIOR FILING DATE: 1999-02-04
; PRIOR APPLICATION NUMBER: 09/122,449
; PRIOR FILING DATE: 1998-07-24
; PRIOR APPLICATION NUMBER: 09/118,205

; PRIOR FILING DATE: 1998-07-16
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 39
; LENGTH: 465
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-092-063-39

Query Match 81.7%; Score 1837.5; DB 14; Length 465;
Best Local Similarity 84.0%; Pred. No. 3.5e-174;
Matches 353; Conservative 25; Mismatches 39; Indels 3; Gaps 3;

Qy 1 MATSGTGFVFMVSVCSAVSHRNQQTWTFEGIFLSSMCPINVSASTLYGIMFDAGSTGT 60
Db 1 MATSGAV-FMLIIACVSGSTVFYREQQTWTFEGVFLSSMCPINVSAGTFYGMFDAGSTGT 59
61 RIHYVTFVQKMPGQLPILLEGVFSVSKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 120
Db 60 RIHYVTFVQKTAGQLPFLGFIPIFSVKPGLSAFVDQPKQGAETVQGLLEVAKDSIPRSHW 119
121 KKTVPVLKATAGRLLPKHAALLPEVKEIFRSPFLVPKGSVIMDGSDEGILAWTV 180
Db 120 ERTPVWLKATAGRLLPKHAALLPEVKEIFRSPFLVPKGSVIMDGSDEGILAWTV 179
181 NFLTQGLHGRQETVGTDLGGASTQTITFLPQPEKTLTQTPRGYLTSEFMFNSYKLYTH 240
Db 180 NFLTQGLHGRQETVGTDLGGASTQTITFLPQPEKTLTQTPRGYLTSEFMFNSYKLYTH 239
241 SYLGFGKLAARLATLGALETEGTDGHTFRSACLPRMLEAEWIFGVVYQYGGNOEGEVSF 300
Db 240 SYLGFGKLAARLATLGALETEGTDGHTFRSACLPRMLEAEWIFGVVYQYGGNOEGEVSF 299
301 EPCYAEVLVRVVRGKLHQPEEVQSGSFYAFSYYVDRAVDTMDIDYKGGILKVEDFERKAR 360
Db 300 EPCYAEVLVRVVRGKLHQPEEVQSGSFYAFSYYVDRAVDTMDIDYKGGILKVEDFERKAR 359
361 EVCNLENFTSGSPFLCNDLSYITALLKDGFGFADSTVLQITKKVNNIETGVALGATFHL 420
Db 360 EVCNLDGSGSGSPFLCNDLTYITALLKDGFGFADSTVLQITKKVNNIETGVALGATFHL 417

RESULT 15
US-09-925-299-876
; Sequence 876, Application US/09925299
; Patent No. US20020055627A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA102
; CURRENT APPLICATION NUMBER: US/09/925,299
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05883
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1556
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 876
; LENGTH: 330
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (97)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (106)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (124)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE

LOCATION: (138)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
NAME/KEY: SITE  
LOCATION: (174)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
NAME/KEY: SITE  
LOCATION: (178)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
NAME/KEY: SITE  
LOCATION: (194)  
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-325-299-876

Query Match 73.8%; Score 1660; DB 9; Length 330;  
Best Local Similarity 97.3%; Pred. No. 1e-156;  
Matches 319; Conservative 0; Mismatches 9; Indels 0; Gaps 0;  
Qy 101 AETVQGLLEVAKDSIPRSHWKTPVVLKATAGLRLLEPHKAKALLFEVKEIFRKSPFLVP 160  
Db 3 ARAVQGLLEVAKDSIPRSHWKTPVVLKATAGLRLLEPHKAKALLFEVKEIFRKSPFLVP 62  
Qy 161 KGSVIMDGSDEGILAWVTNFLTQGLHGRQETVGTLDLGGASTQITFLPQFEKLEQT 220  
Db 63 KGSVIMDGSDEGILAWVTNFLTQGLHGRQETVGTLDLGGASTQITFLPQFEKLEQT 122  
Qy 221 PRGYLTSFEMFNSTYKLYTHSYLGFGLKAARLATLGALETEGTGHTFRSACLPRMLEAE 280  
Db 123 PXGYLTSFEMFNSTYKLYTHSYLGFGLKAARLATLGALETEGTGHTFRSACLPRMLEAE 182  
Qy 281 WIFGGVKYQYGGNOEGVEGPEPCVAEVLVRVGRKLHQPVEVQSGSFYAFSYYVDRAVDT 340  
Db 183 WIFGGVKYQYCXNOEGVEGPEPCVAEVLVRVGRKLHQPVEVQSGSFYAFSYYVDRAVDT 242  
Qy 341 MIDYKGGILKVEDFERKAREVCNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQ 400  
Db 243 MIDYKGGILKVEDFERKAREVCNLENFTSGSPFLCMDLSYITALLKDGFGFADSTVLQ 302  
Qy 401 LTKKVNNIETGVALGATFHLQLSGLISH 428  
Db 303 LTKKVNNIETGVALGATFHLQLSGLISH 330

Search completed: January 29, 2004, 10:14:49  
Job time : 30.5 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 09:35:09 ; Search time 9.5 seconds  
(without alignments)  
2118.677 Million cell updates/sec

Title: US-10-091-085-7

Perfect score: 2252

Sequence: 1 MATSWGTFVFLVSCVSA.....ETGALGATFLLQLSLGISH 428

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_41.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2235	99.2	428	1 ENP5_HUMAN	O75356 homo sapien
2	1979.5	87.9	427	1 ENP5_MOUSE	Q9wuz9 mus musculus
3	1918.5	85.2	469	1 ENP5_MESAU	Q9qyc8 mesocricetu
4	996	44.2	484	1 ENP6_HUMAN	O75354 mus sapien
5	983	43.7	455	1 ENP6_RAT	Q9er31 rattus norv
6	488.5	21.7	518	1 GDAL_YEAST	P32621 saccharomyc
7	484	21.5	455	1 NTPA_PEA	P52914 pisum sativ
8	441.5	19.6	454	1 APY_SOLTU	P80595 solanum tub
9	394	17.5	493	1 ENP1_CHICK	O93295 gallus gall
10	387	17.2	552	1 YBU4_CABEL	O21815 caenorhabdi
11	385	17.1	510	1 ENP1_MOUSE	P55772 mus musculus
12	382.5	17.0	485	1 YY4F_CABEL	Q18411 caenorhabdi
13	375	16.7	513	1 ENP1_BOVIN	O18956 bos taurus
14	371	16.5	630	1 YND1_YEAST	P40009 saccharomyc
15	367	16.3	510	1 ENP1_PIG	Q9myu4 sus scrofa
16	366.5	16.3	613	1 ENP4_MOUSE	Q9dbt4 mus musculus
17	365	16.3	510	1 ENP1_HUMAN	P49961 homo sapien
18	363.5	16.1	511	1 ENP1_RAT	P97687 rattus norv
19	363	16.1	616	1 ENP4_HUMAN	Q9y227 homo sapien
20	357.5	15.9	529	1 ENP3_HUMAN	O75355 homo sapien
21	351.5	15.6	495	1 ENP2_HUMAN	Q9y513 homo sapien
22	321.5	14.3	495	1 ENP2_MOUSE	O55026 mus musculus
23	314.5	14.0	495	1 ENP2_RAT	O35795 rattus norv
24	310.5	13.8	494	1 ENP2_CHICK	P79784 gallus gall
25	150	6.7	628	1 NTP1_TOXGO	Q27893 toxoplasma
26	148	6.6	628	1 NTP2_TOXGO	Q27895 toxoplasma
27	141	6.3	592	1 NTP4_TOXGO	P52913 toxoplasma
28	105.5	4.7	553	1 VGI_SPV4	P11333 spiroplasma
29	104.5	4.6	535	1 DEXB_STEPRN	Q54796 streptococc
30	100.5	4.5	634	1 SELB_MOOTH	Q6455 moorella th
31	100.5	4.5	1151	1 Y245_TREPA	O83273 treponema p
32	99	4.4	711	1 DPP3_YEAST	Q08225 saccharomyc
33	98.5	4.4	4427	1 PKSL_BACSU	Q05470 bacillus su

34	96.5	4.3	287	1	RM07_SCHPO	O14337 schizosacch
35	96	4.3	477	1	UR11_DESRO	P98119 desmodus to
36	92	4.1	290	1	HEM6_CAUCR	Q9aat8 caulobacter
37	92	4.1	424	1	Y115_ARCFU	P56969 archaeglob
38	92	4.1	1220	1	PTC1_BRARE	Q98864 brachydanio
39	92	4.1	1549	1	YIR3_YEAST	P40438 saccharomyc
40	92	4.1	1549	1	YJW2_YEAST	P40890 saccharomyc
41	91.5	4.1	796	1	TLR6_HUMAN	Q9y2c9 homo sapien
42	91	4.0	417	1	GLYA_BORBU	O51547 borrelia bu
43	91	4.0	946	1	ITH2_MOUSE	Q61703 mus musculu
44	90.5	4.0	279	1	HEM6_RICPR	Q9zc86 rickettsia
45	90.5	4.0	2076	1	PAS1_YARLI	P34229 y fatty aci

## ALIGNMENTS

RESULT 1	ENP5_HUMAN	STANDARD;	PRT;	428 AA.
ID	ENP5_HUMAN	AC	O75356;	
DT	16-OCT-2001	(Rel. 40, Created)		
DT	16-OCT-2001	(Rel. 40, Last sequence update)		
DT	16-OCT-2001	(Rel. 40, Last annotation update)		
DE	Ectonucleoside triphosphate diphosphohydrolase 5 precursor			
DE	(EC 3.6.1.6) (NTPDase5) (Nucleoside diphosphatase) (CD39 antigen-like			
DE	4) (ER-UDPase).			
GN	ENTPD5 OR CD39L4.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Leukemia;			
RX	MEDLINE=98341119; PubMed=9676430;			
RA	Chadwick B.P., Frischauf A.-M.;			
RT	"The CD39-like gene family: identification of three new human members			
RT	(CD39L2, CD39L3, and CD39L4), their murine homologues, and a member of			
RT	the gene family from Drosophila melanogaster.";			
RL	Genomics 50:357-367(1998).			
CC	-I- FUNCTION: LIKELY TO PROMOTE REGLYCOSYLATION REACTIONS INVOLVED IN			
CC	GLYCOPROTEINS FOLDING AND QUALITY CONTROL IN THE ENDOPLASMIC			
CC	RETICULUM. HYDROLYZES UDP, GDP AND IDP BUT NOT ANY OTHER			
CC	NUCLEOSIDE DI-, MONO- OR TRIPHOSPHATES, NOR THIAMINE PYROPHOSPHATE			
CC	(BY SIMILARITY).			
CC	-I- CATALYTIC ACTIVITY: A nucleoside diphosphate + H(2)O = a			
CC	nucleotide + phosphate.			
CC	-I- COFACTOR: REQUIRES CALCIUM AND MAGNESIUM.			
CC	-I- SUBCELLULAR LOCATION: Endoplasmic reticulum lumen (By similarity).			
CC	-I- TISSUE SPECIFICITY: EXPRESSED IN ADULT LIVER, KIDNEY, PROSTATE,			
CC	TESTIS AND COLON. MUCH WEAKER EXPRESSION IN OTHER TISSUES.			
CC	-I- SIMILARITY: BELONGS TO THE GDAL / CD39 NTPASE FAMILY.			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	EMBL: AF039918; AAC39885.1; ..			
DR	Genew; HGNC:3367; ENTPD5.			
DR	MIM; 603162; ..			
DR	GO; GO:0004002; F:adenosinetriphosphatase activity; TAS.			
DR	InterPro; IPR000407; GDAL_CD39_NTPase.			
DR	Pfam; PF01150; GDAL_CD39; 1.			
DR	PROSITE; PS01238; GDAL_CD39_NTPASE; FALSE NEG.			
KW	Hydrolase; Transmembrane; Glycoprotein; Calcium; Magnesium;			
KW	Endoplasmic reticulum; Signal.			
FT	SIGNAL 1 20 POTENTIAL.			
FT	CHAIN 21 428 ECTONUCLEOSIDE TRIPHOSPHATE			

FT CARBOHYD 42 42 DIPHOSEPHOHYDROLASE 5.  
FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 232 232 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 368 368 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 428 AA; 47517 MW; 830437A155D84DD CRC64;  
Query Match 99.2%; Score 2235; DB 1; Length 428;  
Best Local Similarity 99.3%; Pred. No. 5.2e-172;  
Matches 425; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
Qy 1 MATSGTFFVFMVSVCSAVSHRNQQTWFGIFLSSMCPINVSASTLYGIMFDAGSTGT 60  
Db 1 MATSGTFFVFMVSVCSAVSHRNQQTWFGIFLSSMCPINVSASTLYGIMFDAGSTGT 60  
Qy 61 RIHYTVFQKMPGOLPILGEVFDVSRPGLSAFVDQPKQAGTVOGLEVAKDSIPRSHW 120  
Db 61 RIHYTVFQKMPGOLPILGEVFDVSRPGLSAFVDQPKQAGTVOGLEVAKDSIPRSHW 120  
Qy 121 KKTVPVLCATAGLRLLPEHKAKALLFEVKEIFRKSPLVPKGSVSMITQDEGIFAWTV 180  
Db 121 KKTVPVLCATAGLRLLPEHKAKALLFEVKEIFRKSPLVPKGSVSMITQDEGIFAWTV 180  
Qy 181 NFLTGOLGHRQETVGTLDGASTQITFLPQPKTLEOTPRGYLTSFEMFNSTYKLYTH 240  
Db 181 NFLTGOLGHRQETVGTLDGASTQITFLPQPKTLEOTPRGYLTSFEMFNSTYKLYTH 240  
Qy 241 SYLGFGLKAARLATLGALTEGTGHTFRSACLPRWLEAEWIFGVYQVYGGNQEVEGF 300  
Db 241 SYLGFGLKAARLATLGALTEGTGHTFRSACLPRWLEAEWIFGVYQVYGGNQEVEGF 300  
Qy 301 EPCVAELVRVVRGKLHQPVEVQSGSFYAFSYVDVTDMDIDYKGGILKVEDFERKAR 360  
Db 301 EPCVAELVRVVRGKLHQPVEVQSGSFYAFSYVDVTDMDIDYKGGILKVEDFERKAR 360  
Qy 361 EVCNLENFTSGSPFLCNDLSYITALLKDGFGFADSTVLQTKVNNIETGALGATFHL 420  
Db 361 EVCNLENFTSGSPFLCNDLSYITALLKDGFGFADSTVLQTKVNNIETGALGATFHL 420  
Qy 421 LOSLGISH 428  
Db 421 LOSLGISH 428

## RESULT 2

ENPS\_MOUSE STANDARD; PRT; 427 AA.  
AC Q9WU29, 070214;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Ectonucleoside triphosphate diphosphohydrolase 5 precursor  
DE (EC 3.6.1.6) (NTPDase5) (Nucleoside diphosphatase) (CD39 antigen-like  
DE 4) (ER-UPDase).  
GN ENTPD5 OR CD39L4.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID:10090;  
RN [1]  
SEQUENCE FROM N.A., SUBCELLULAR LOCATION, AND CHARACTERIZATION.  
RC TISSUE=Liver;  
RC MEDLINE=99298181; PubMed=10369669;  
RA Trombetta E.S., Helenius A.;  
RT "CDNA cloning and chromosomal mapping of a mouse gene with homology to  
RT NTPases.",  
RL Mamm. Genome 9:162-164 (1998).  
RN [2]  
SEQUENCE FROM N.A., SUBCELLULAR LOCATION, AND CHARACTERIZATION.  
RC TISSUE=Liver;  
RC MEDLINE=99298181; PubMed=10369669;  
RA Trombetta E.S., Helenius A.;  
RT "Glycoprotein regucosylation and nucleotide sugar utilization in the  
RT secretory pathway: identification of a nucleoside diphosphatase in the

endoplasmic reticulum.",  
EMBO J. 18:3282-3292 (1999).  
[3]  
SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=Kidney;  
RX MEDLINE=21085660; PubMed=11217851;  
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,  
RA Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,  
RA Schram L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
RA Gustinich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
RA Suzuki H., Sato K., Schoenbach C., Seva T., Shibata Y., Storch K.-F.,  
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,  
RA Hayaishizaki Y.;  
RT "Functional annotation of a full-length mouse cDNA collection.",  
RL Nature 409:685-690 (2001).  
CC -1- FUNCTION: LIKELY TO PROMOTE REGLYCOSYLATION REACTIONS INVOLVED IN  
CC GLYCOPROTEIN FOLDING AND QUALITY CONTROL IN THE ENDOPLASMIC  
CC RETICULUM. HYDROLYZES UDP, GDP AND IDP BUT NOT ANY OTHER  
CC NUCLEOSIDE DI-, MONO- OR TRIPHOSPHATES, NOR THIAMINE  
CC PYROPHOSPHATE.  
CC -1- CATALYTIC ACTIVITY: A nucleoside diphosphate + H(2)O = a  
CC nucleotide + phosphate.  
CC -1- COFACTOR: REQUIRES CALCIUM AND MAGNESIUM.  
CC -1- SUBCELLULAR LOCATION: Endoplasmic reticulum lumen.  
CC -1- TISSUE SPECIFICITY: Ubiquitous.  
CC -1- PTM: GLYCOSYLATED WITH HIGH MANNOSE N-LINKED GLYCANS.  
CC -1- MISCELLANEOUS: OPTIMAL PH IS NEUTRAL.  
CC -1- SIMILARITY: BELONGS TO THE GDAL / CD39 NTPASE FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC -----  
CC EMBL; AF006482; AAC05181.1; -;  
CC EMBL; AJ238636; CAB45533.1; -;  
CC EMBL; AK002618; BAB22234.1; -;  
CC MGD; MGI:1321385; Entpd5  
CC InterPro; IPR000407; GDAL\_CD39\_NTPase.  
CC Pfam; PFO1150; GDAL\_CD39; 1.  
CC PROSITE; PS01238; GDAL\_CD39\_NTPASE; FALSE NEG.  
KW Hydrolase; Transmembrane; Glycoprotein; Calcium; Magnesium;  
KW Endoplasmic reticulum; Signal.  
FT SIGNAL 1 18 POTENTIAL.  
FT CHAIN 19 427 ECTONUCLEOSIDE TRIPHOSPHATE  
FT FT DIPOHOSPHOHYDROLASE 5.  
FT FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT FT CARBOHYD 41 41 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT FT CARBOHYD 231 231 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT FT CONFLICT 390 390 F -> L (IN REF. 1).  
FT FT CONFLICT 394 427 DGTLLTQTKVNNIETGALGATFHLQSLGITS -> ERH  
FT FT PLTAHKESEQRHQRWLGGLGHLSPAPVSGHQLRPSSTSEAC  
FT FT ISEPFVFSQGVDSSTFSLDGKAWPETR (IN REF. 1).  
SQ SEQUENCE 427 AA; 47101 MW; 2F9DA2C342C55577 CRC64;  
Query Match 87.9%; Score 1979.5; DB 1; Length 427;  
Best Local Similarity 87.6%; Pred. No. 1.7e-151;  
Matches 374; Conservative 24; Mismatches 28; Indels 1; Gaps 1;  
Qy 1 MATSWCTVFFMLVSVCSAVSHRNQQTWFGIFLSSMCPINVSASTLYGIMFDAGSTGT 60

```

Db 1 MATSWGAV-FMLIIACVSGSTVFYREQQTWPEGVFLSSMCPINVSAGTFYGIIMFDAGSTGT 59
Qy 61 RIHVYTFVQKMPQLPTEGEVDSVKPGLSAFVDPKQGAETVQGLLEVAKDSIPRSHW 120
Db 60 RIHVYTFVQKMPQLPTEGEVDSVKPGLSAFVDPKQGAETVQGLLEVAKDSIPRSHW 119
Qy 121 KPTPVVLKATAGLRLPEHAKALLFEVKEIFRKSPFLVPKGSVIMTQDEGIFAWTV 180
Db 120 ERTFVVLKATAGLRLPEHAKALLFEVKEIFRKSPFLVPKGSVIMTQDEGIFAWTV 179
Qy 181 NFLTGQHGHRQETVGTDLGGASTQITFLPQPEKTELETPRGYLSFEMFNSTYKLYTH 240
Db 180 NFLTGQHGHRQETVGTDLGGASTQITFLPQPEKTELETPRGYLSFEMFNSTYKLYTH 239
Qy 241 SYLGFGKKAARLATLGALETGDTGHTFRSACLPRLWEAEWIFGGVKYQYGGNQEVEVGF 300
Db 240 SYLGFGKKAARLATLGALETGDTGHTFRSACLPRLWEAEWIFGGVKYQYGGNQEVEVGF 299
Qy 301 EPCYAEVLRVVRGKLHQPVEVQSGSFYAFSYYYDRAVDTMDIDYKGGILKVEDFERKAR 360
Db 300 EPCYAEVLRVVRGKLHQPVEVQSGSFYAFSYYYDRAVDTMDIDYKGGILKVEDFERKAR 359
Qy 361 EVCNLENFTSGSPFLCMDLSYITALLKQGFADSTVLQTKKVNNIETGHALGATFHL 420
Db 360 EVCNLENFTSGSPFLCMDLSYITALLKQGFADSTVLQTKKVNNIETGHALGATFHL 419
Qy 421 LQSLGTS 427
Db 420 LQSLGTS 426

RESULT 3
ENP6_HUMAN STANDARD; PRT; 469 AA.
ID ENP6_HUMAN STANDARD; PRT; 469 AA.
AC Q90YCB; Q90YCB; 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ectonucleoside triphosphate diphosphohydrolase 5 precursor
DE (EC 3.6.1.6) (NTPDase5) (Nucleoside diphosphatase) (CD39 antigen-like
DE (3R-UDPase) (Proto-oncogene cph).
GN NTPD5 OR CD39L4 OR CPH.
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
OX NCBI_TaxID=10036;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99142925; PubMed=9989819;
RA Velasco J.A., Avila M.A., Notario V.;
RT "The product of the cph oncogene is a truncated, nucleotide-binding
RT protein that enhances cellular survival to stress."
RL Oncogene 18:689-701(1999).
CC -1- FUNCTION: LIKELY TO PROMOTE REGLYCOSYLATION REACTIONS INVOLVED IN
CC GLYCOPROTEIN FOLDING AND QUALITY CONTROL IN THE ENDOPLASMIC
CC RETICULUM. HYDROLYZES UDP, GDP AND IDP BUT NOT ANY OTHER
CC NUCLEOSIDE DI-, MONO- OR TRIPHOSPHATES, NOR THIAMINE PYROPHOSPHATE
CC (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: A nucleoside diphosphate + H(2O) = a
CC nucleotide + phosphate.
CC -1- COFACTOR: REQUIRES CALCIUM AND MAGNESIUM (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Endoplasmic reticulum lumen (By similarity).
CC -1- TISSUE SPECIFICITY: EXPRESSED IN FETAL CELLS AND MOST ADULT
CC TISSUES.
CC -1- DISEASE: CPH ONCOGENE HAS TRANSFORMING CAPACITY AND TUMORIGENIC
CC POTENTIAL.
CC -1- SIMILARITY: BELONGS TO THE GDAL / CD39 NTPASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements/
CC or send an email to license@ebi.ac.uk).
CC
CC EMBL; AF084568; AAF22931.1; ALT_TERM.
DR EMBL; AF084569; AAF22932.1; -.
DR InterPro; IPR000407; GDAL_CD39_NTPase.
DR Pfam; PF01150; GDAL_CD39; 1.
DR PROSITE; PS01238; GDAL_CD39_NTPASE; FALSE NEG.
KW Hydrolase; Transmembrane; Glycoprotein; Calcium; Magnesium;
KW Endoplasmic reticulum; Signal; Proto-oncogene.
FT SIGNAL 1 18
FT CHAIN 19 469
FT CARBOHYD 42 42
FT CARBOHYD 232 232
FT SEQUENCE 469 AA; 52125 MW; 03D8A23E0C73474B CRC64;
SQ
Query Match 85.2%; Score 1918.5; DB 1; Length 469;
Best Local Similarity 86.0%; Pred. No. 1.6e-146;
Matches 361; Conservative 22; Mismatches 34; Indels 3; Gaps 1;
Qy 1 MATSWGAV-FMLIIACVSGSTVFYREQQTWPEGVFLSSMCPINVSAGTFYGIIMFDAGSTGT 60
Db 1 MATSWGAV-FMLIIACVSGSTVFYREQQTWPEGVFLSSMCPINVSAGTFYGIIMFDAGSTGT 60
Qy 61 RIHVYTFVQKMPQLPTEGEVDSVKPGLSAFVDPKQGAETVQGLLEVAKDSIPRSHW 120
Db 61 RIHVYTFVQKMPQLPTEGEVDSVKPGLSAFVDPKQGAETVQGLLEVAKDSIPRSHW 120
Qy 121 KPTPVVLKATAGLRLPEHAKALLFEVKEIFRKSPFLVPKGSVIMTQDEGIFAWTV 180
Db 121 KPTPVVLKATAGLRLPEHAKALLFEVKEIFRKSPFLVPKGSVIMTQDEGIFAWTV 180
Qy 181 NFLTGQHGHRQETVGTDLGGASTQITFLPQPEKTELETPRGYLSFEMFNSTYKLYTH 240
Db 181 NFLTGQHGHRQETVGTDLGGASTQITFLPQPEKTELETPRGYLSFEMFNSTYKLYTH 240
Qy 241 SYLGFGKKAARLATLGALETGDTGHTFRSACLPRLWEAEWIFGGVKYQYGGNQEVEVGF 300
Db 241 SYLGFGKKAARLATLGALETGDTGHTFRSACLPRLWEAEWIFGGVKYQYGGNQEVEVGF 300
Qy 301 EPCYAEVLRVVRGKLHQPVEVQSGSFYAFSYYYDRAVDTMDIDYKGGILKVEDFERKAR 360
Db 301 EPCYAEVLRVVRGKLHQPVEVQSGSFYAFSYYYDRAVDTMDIDYKGGILKVEDFERKAR 360
Qy 361 EVCNLENFTSGSPFLCMDLSYITALLKQGFADSTVLQTKKVNNIETGHALGATFHL 420
Db 361 EVCNLENFTSGSPFLCMDLSYITALLKQGFADSTVLQTKKVNNIETGHALGATFHL 420
RESULT 4
ENP6_HUMAN STANDARD; PRT; 484 AA.
ID ENP6_HUMAN STANDARD; PRT; 484 AA.
AC Q75354; Q9UJD1; 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ectonucleoside triphosphate diphosphohydrolase 6 (EC 3.6.1.6)
DE (NTPDase6) (CD39 antigen-like 2).
GN NTPD6 OR CD39L2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98341119; PubMed=9676430;
RA Chadwick B.P., Frieschaut A.-M.;
RT "The CD39-like gene family: identification of three new human members
RT (CD39L2, CD39L3, and CD39L4), their murine homologues, and a member of

```



CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.ebi.ac.uk/announcement/>  
 CC or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).

DR EMBL; AJ277748; CAC16598.1; --  
 DR InterPro; IPR000407; GDAI\_CD39\_NTPase.  
 DR Pfam; PF01150; GDAI\_CD39; 1.  
 DR PROSITE; PS01238; GDAI\_CD39\_NTPase; FALSE NEG.  
 KW Hydrolase; Transmembrane; Glycoprotein; Calcium; Magnesium;  
 KW Signal-anchor; Golgi stack.  
 FT DOMAIN 1 12 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 13 32 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)  
 FT (POTENTIAL).  
 FT DOMAIN 33 455 LUMENAL (POTENTIAL).  
 FT CARBOHYD 192 192 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 256 256 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 455 AA; 49899 MW; 19A22E8BAEF0F77B CRC64;

Query Match 43.7%; Score 983; DB 1; Length 455;  
 Best Local Similarity 54.0%; Pred. No. 2.1e-71;  
 Matches 204; Conservative 49; Mismatches 119; Indels 6; Gaps 5;

QY 49 YGIMFDAGSTGTRIHYVTFVQKMPGQLPILGEVDSVKPGLSAFVDQPKQGAETVQGL 108  
 DB 74 YGIMFDAGSTGTRIHYVTFVQKMPGQLPILGEVDSVKPGLSAFVDQPKQGAETVQGL 132  
 QY 109 EVAKDSIPRSHWKTTPVVLKATAGLLPEHKAKALLFEVKEIFRSPFLVPGSVSINT 168  
 DB 133 NVAKQHPIYDFWKATPLVLKATAGLLPEHKAKALLFEVKEIFRSPFLVPGSVSINT 192  
 QY 169 GODEGIFAWTVNFTLQGLHGRQETVGTDLGASTQITFLPQFETLROTGRVLTSP 228  
 DB 193 GTDEGVSAITVNTLQGLHGRQETVGTDLGASTQITFLPQFETLROTGRVLTSP 252  
 QY 229 EMPNSTYKLYTHSYLFGGLKAARLATLGALETE-GTDGHTFRSACLPRLWEAEWIFGGVK 287  
 DB 253 QMFNRTFKLYSYLFGGLKAARLATLGALETE-GTDGHTFRSACLPRLWEAEWIFGGVK 312  
 QY 288 YQGGNQEGRVG-FPCYARVLAVRGKLHQPEVQSGFYAFSYYYDRAVDTMDIDYEK 346  
 DB 313 YRISGQK--AVGAYELCASRVSEVLKRVKVRTEAQAQHVDFYAFSYYYDRAVDTMDIDYEK 370  
 QY 347 GGLIKVEDFERKAREVCNLENFTSGSPFLCMLDSYITALLKDGFGFADSTVLQITKVN 406  
 DB 371 GGSVLVGDFFIAKYVCTLETOPSPFPACMDLTYSLLLHE-FGPPGDKVLKARKID 429  
 QY 407 NIETGVALGATPHLLQSL 424  
 DB 430 NVETSWALGAIFHYDLSL 447

RESULT 6  
 GDAI\_YEAST  
 ID GDAI\_YEAST STANDARD; PRT; 518 AA.  
 AC P32621;  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DT 01-OCT-1993 (Rel. 27, Last sequence update)  
 DT 15-SEP-2003 (Rel. 42, Last annotation update)  
 DE Guanosine-diphosphatase (EC 3.6.1.42) (GDPase).  
 GN GDAI OR YEL042W OR SYGP-ORF16.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RC SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RC STRAIN=G2-9;  
 RX MEDLINE=93308137; PubMed=8391537;  
 RA Abeijon C., Yanagisawa K., Mandon E.C., Haessler A., Moremen K.,  
 RA Hirschberg C.B., Robbins P.W.;

"Guanosine diphosphatase is required for protein and sphingolipid  
 glycosylation in the Golgi lumen of *Saccharomyces cerevisiae*.";  
 J. Cell Biol. 122:307-323(1993).

[2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=S288C / AB972;  
 RX PubMed=9169868;  
 RA Dietrich F.S., Mulligan J.T., Hennessey K.M., Yelton M.A., Allen E.,  
 RA Araujo R., Aviles E., Berno A., Brennan T., Carpenter J., Chen E.,  
 RA Cherry J.M., Chung E., Duncan M., Guzman E., Hartzell G.,  
 RA Hunnicke-Smith S., Hyman R.W., Kayser A., Komp C., Lashkari D., Lew H.,  
 RA Lin D., Mosedale D., Nakahara K., Namath A., Norgren R., Oefner P.,  
 RA Oh C., Petel F.X., Roberts D., Sehl P., Schramm S., Shogren T.,  
 RA Smith V., Taylor P., Wei Y., Botstein D., Davis R.W.;  
 RT "The nucleotide sequence of *Saccharomyces cerevisiae* chromosome V.";  
 RL Nature 387:78-81(1997).  
 CC -1- FUNCTION: AFTER TRANSFER OF SUGARS TO ENDOGENOUS MACROMOLECULAR  
 CC ACCEPTORS, THE ENZYME CONVERTS NUCLEOSIDE DIPHOSPHATES TO  
 CC NUCLEOSIDE MONOPHOSPHATES WHICH IN TURN EXIT THE GOLGI LUMEN IN  
 CC A COUPLED ANTIporter REACTION, ALLOWING ENTRY OF ADDITIONAL  
 CC NUCLEOTIDE SUGAR FROM THE CYTOSOL.  
 CC -1- CATALYTIC ACTIVITY: GDP + H(2)O = GMP + phosphate.  
 CC -1- PATHWAY: Glycosylation.  
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein. Golgi.  
 CC -1- SIMILARITY: BELONGS TO THE GDAI / CD39 NTPASE FAMILY.

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.ebi.ac.uk/announcement/>  
 CC or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).

EMBL; L19560; AAA34656.1; --  
 EMBL; U18779; AAB5000.1; --  
 EIR; A40732; A40732.  
 DR SGD; S0000768; GDAI.  
 DR GO; GO:0004382; F:guanosine diphosphatase activity; IDA.  
 DR GO; GO:0045134; F:uridine diphosphatase activity; IDA.  
 DR InterPro; IPR000407; GDAI\_CD39\_NTPase.  
 DR Pfam; PF01150; GDAI\_CD39; 1.  
 DR PROSITE; PS01238; GDAI\_CD39\_NTPase; 1.  
 KW Hydrolase; Golgi stack; Glycoprotein; Transmembrane; Signal-anchor.  
 FT DOMAIN 1 9 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 10 24 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)  
 FT (POTENTIAL).  
 FT DOMAIN 25 518 LUMENAL (POTENTIAL).  
 FT CARBOHYD 41 41 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 280 280 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 335 335 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 518 AA; 56821 MW; 9A61182D5ED22ADF CRC64;

Query Match 21.7%; Score 488.5; DB 1; Length 518;  
 Best Local Similarity 33.0%; Pred. No. 1.3e-31;  
 Matches 143; Conservative 56; Mismatches 159; Indels 75; Gaps 15;

QY 49 YGIMFDAGSTGTRIHYVTFVQKMPGQLPILGEVDSVKPGLSAFVDQPKQGAETVQGL 107  
 DB 93 YVIMFDAGSTGSRVHYIKFDVCTSP---PTLLDEKFDMLPEGLSSFDTSVGAANSLDPL 149  
 QY 108 LEVAKDSIPRSHWKTTPVVLKATAGLLPEHKAKALLFEVKEIFRSPFLVPGK-SVS 165  
 DB 150 LKVMYVVIKARSCPTFAVAVKATAGLLGDAKSSKILSAVRDHLKDYFPVVEGDDVS 209  
 QY 166 IMTGDEGIFAWTVNFTLQGL--HGHRQETVGTDLGASTQITFLPQFETLROTGRV 223  
 DB 210 IMGDEEGVFAWITTYLLGNIGANGPKLPATAVFDLGGGSGTQIFPEPTPIKEKMDGE 269  
 QY 224 YLTSFEMFNSTYKLYTHSYLFGGLKAAR-----LATLGALETEGTDGHTFRSAC 272  
 DB 270 HKFDLKFQDENTYLYQFSLHGYGLKEGRKNVSVLVENALKQGLKDKNTKTHQLSSPC 329

QY 273 LPRWLEA-EWI-----FGVKYQYGG-----NQEVEGFPCYAEVLR 309  
 Db 330 LPPKVNATNEKVTLESKETVTIDFGDEPSGAQCRFLTDEILNKDAQQCPSCF----- 385  
 QY 310 VVRGKLHOPEV-----QRGSFYAFSYYYDRAVDTDM-IDYEKGILKVEDFERKAREVC- 363  
 Db 386 ---NGVHQPSLVRTFKESNDIYIFSYFYDRTRPLGMPISF-----TLNELNDLARIYCK 436  
 QY 364 -----DNLENTSGPFLCMLDSYITALLKDGFGFADSVLQITKKVNIET 410  
 Db 437 GEETWNSVFGSIAGSLDELSDSHF-CLDSFQVSLLLHTGYDIPLOBELRTGKIANKEI 495  
 QY 411 GWALGATFHLLQS 423  
 Db 496 GWCLGASLPLKA 508

RESULT 7  
 NTPA\_PEA STANDARD; PRT; 455 AA.  
 AC P52914;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Nucleoside-triphosphatase (EC 3.6.1.15) (Nucleoside triphosphate  
 DE phosphohydrolase) (NTPase) (Apyrase).  
 OS Pisum sativum (Garden pea).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosoids I; Fabales; Fabaceae; Papilionoideae; Viciaeae; Pisum.  
 OX NCBI\_TaxID=3888;  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RC STRAIN=cv. Alaska; TISSUE=Plumule;  
 RX MEDLINE=96197404; PubMed=8615230;  
 RA Hsieh H., Tong C.G., Thomas C., Roux S.J.;  
 RT "Light-modulated abundance of an mRNA encoding a calmodulin-regulated,  
 RT chromatin-associated NTPase in pea";  
 RL Plant Mol. Biol. 30:135-147(1996).  
 RN [2]  
 RC SEQUENCE FROM N.A.  
 RC STRAIN=cv. Alaska; TISSUE=Stem;  
 RA Shibata K., Abe S., Davies E.;  
 RT "Structure of the coding region and mRNA variants of the apyrase from  
 RT Pisum sativum";  
 RL Acta Physiol. Plant. 20:3-13(2001).  
 CC -1- FUNCTION: MIGHT BE INVOLVED IN RNA TRANSPORT OUT OF NUCLEI.  
 CC -1- CATALYTIC ACTIVITY: NTP + H(2)O = NDP + phosphate.  
 CC -1- SUBCELLULAR LOCATION: Nuclear.  
 CC -1- SIMILARITY: BELONGS TO THE GDAL / CD39 NTPASE FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 DR EMBL; Z32743; CAA83655.1; -  
 DR EMBL; AB027614; BAB18900.1; -  
 DR PIR; S65147; S48859.  
 DR InterPro; IPR000407; GDAL\_CD39\_NTPase.  
 DR Pfam; PF011150; GDAL\_CD39\_1.  
 DR PROSITE; PS01239; GDAL\_CD39\_NTPASE; 1.  
 KW Hydrolyase; Nuclear protein.  
 SQ SEQUENCE 455 AA; 50072 MW; 50FDF0023ABC4299 CRC64;  
 Query Match 21.5%; Score 484; DB 1; Length 455;  
 Best Local Similarity 33.4%; Pred. No. 2.5e-31;  
 Matches 139; Conservative 62; Mismatches 161; Indels 54; Gaps 14;  
 QY 49 YGIMFDAGSTGTRIHYVTFVQKMPGQLPILEG-EVFDSVKPGLSAFVDQPKQAGTVOGL 107

Db 44 YAVVFDAGSTGSRHYYHFNQNL-DLLHIGKGYEYNNKITPGSSYANNPEQAASLIPL 102  
 QY 108 LEVAKDSISRSHWKTTPVVLKATAGURLLPEHKAKALLFEVKEIF-RKSPFLPKGSVSI 166  
 Db 103 LEQADVDDDDQPKTPVRLGATAGURLLNGDASEKILQSVRDLNRSRTFNQPDVSI 162  
 QY 167 MTQDDEGIFAWVTNFLTQGLHGRQETVGTLDLGGASTQITFLPQEFKLEQTPR----- 222  
 Db 163 IDGTQSSYLWVTNVALGNLKKYTKTVGVIDLGGSVQMAVAVS-KKTAKNAPKAVDG 221  
 QY 223 --GYLTSFEMFNSTYKLYTHSYLGFGLKAARLATLGALETEGTDGHTFRS--ACLPRWLE 278  
 Db 222 DDPIYIKVVLKGIPIYDLVHSHYLFHFGREASRABILKL-----TPRSPNCLLAGFN 272  
 QY 279 AEWIFGGVKYQYGGNQEVEGFPCYAEVLRVVRGKLHOPEEVEQVGSF----- 326  
 Db 273 GIYTSGEERFKATAYTSG-ANFNKCKNTIRKAL--KLNTPCPYQNTCGFIWNGGGNGQ 329  
 QY 327 ---YAFSYYYDRAVDTDMIDYKGG-ILKVEDFERKAREVCD-NLENFTSGSPFL----- 376  
 Db 330 KNLFASSSFYLPEDITGMVDASTPNEFILRPVDIETKAKEACALNFEDAKSTYFFLDKKNV 389  
 QY 377 ---CMDLSYITALLKDGFGFADSVLQITKKVNN-----IETGVALGATPHLLQSL 424  
 Db 390 ASYVCMDLIYQVYLLVDFGLDPLQKITGSKETIEYQDAIVEAAWPLGNAVEAISAL 445

RESULT 8  
 APY\_SOLTU STANDARD; PRT; 454 AA.  
 ID AC P80595; Q43164;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE Apyrase precursor (EC 3.6.1.5) (ATP-diphosphatase) (Adenosine  
 DE diphosphatase) (ADPase) (ATP-diphosphohydrolase).  
 GN RROPI.  
 OS Solanum tuberosum (Potato).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 OC Asteridae; Lamiales; Solanales; Solanaceae; Solanum.  
 OX NCBI\_TaxID=4113;  
 RN [1]  
 RC SEQUENCE FROM N.A., AND SEQUENCE OF 59-160; 236-253 AND 332-345.  
 RC TISSUE=Tuber;  
 RX MEDLINE=96158985; PubMed=8579614;  
 RA Handa M., Guidotti G.;  
 RT "Purification and cloning of a soluble ATP-diphosphohydrolase  
 RT (apyrase) from potato tubers (Solanum tuberosum).";  
 RL Biochem. Biophys. Res. Commun. 218:916-923(1996).  
 RN [2]  
 RC SEQUENCE OF 42-54; 68-95 AND 236-253.  
 RC STRAIN=cv. Desiree;  
 RX MEDLINE=96355615; PubMed=8703025;  
 RA Vasconcelos E.G., Ferreira S.T., de Carvalho T.M.U., de Souza W.,  
 RA Kettlun A.M., Mancilla M., Valenzuela M.A., Vertovski-Almeida S.;  
 RT "Partial purification and immunohistochemical localization of ATP  
 RT diphosphohydrolase from Schistosoma mansoni. Immunological cross-  
 RT reactivities with potato apyrase and Toxoplasma gondii nucleoside  
 RT triphosphate hydrolase";  
 RL J. Biol. Chem. 271:22139-22145(1996).  
 CC -1- FUNCTION: CATALYZES THE HYDROLYSIS OF PHOSPHOANHYDRIDE BONDS OF  
 CC NUCLEOSIDE TRI- AND DI-PHOSPHATES.  
 CC -1- CATALYTIC ACTIVITY: ATP + 2 H(2)O = AMP + 2 phosphate.  
 CC -1- COFACTOR: CALCIUM.  
 CC -1- SUBCELLULAR LOCATION: Membrane-associated (Probable).  
 CC -1- PTM: THE N-TERMINUS IS BLOCKED.  
 CC -1- SIMILARITY: BELONGS TO THE GDAL / CD39 NTPASE FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its





```

123 TAGMRLLEQNSTKAEQVFAEVSRAIRFPP--VDFRGAQILTNGEEGSGFQWITWYLLLET 180
187 L-----HGHROETVGLDGGASTQITTELPOFEKTELEQPRGLVTSFEMFNSTYK 236
181 LIKESFAGKEHPONTEVLGALDGGASTQITFGV--TIEDKNTSVL--FRUYGNYIS 236
237 LYTHSYLGFGLKAARLATLGALETGDTGHTFRACLPRLWEAEWIFGGVKYQYGGNQEE- 295
237 LYTHSYLCYQIQASKRLMALHODGSVQNI SHPCYPK-----GYRRI 280
296 ---GEVGFPCY-----AEVURVVR-----GKHLQ 317
281 ITIAEIVDSPCVPTFMSLSPAQLTGTGTPNPAACPTAILKFLNLTCGANRTCGFDGVYQ 340
318 PEEVQSGSFVAFS--YYVDRA-----VDTMDIV--EKGILKVEDFERKAREVC 363
341 PP--VRGQFFAFAGFYITFTSLNLTGQOSLSHVNATVMDFCNKWSSELVETFPNK----- 394
364 DNLENFTSGSPFLCWDLSYITALLKDGFGFADST--VLQLTKKVNNIETGWLGAATPHL 420
395 EHLHTY-----CVVGLYIILTLAVDGVKFDHTWSNIHFSQKAGNADIGWTLGLFMLNL 446

RESULT 10
YBU4 CAEEL STANDARD; PRT; 552 AA.
AC Q21815;
DT 01-NOV-1997 (Rel. 35, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical protein R07E4.4 in chromosome X.
GN R07E4.4.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Miller N.;
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
REVISIONS.
RA Waterston R.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: BELONGS TO THE GDA1 / CD39 NTPASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U39652; AAA0403.2;
DR WormPep; R07E4.4; CE28748.
DR InterPro; IPR000407; GDA1_CD39_NTPase.
DR Pfam; PF01150; GDA1_CD39; 1.
DR PROSITE; PS01238; GDA1_CD39_NTPASE; 1.
KW Hypothetical protein; Transmembrane; Hydrolase.
FT TRANSMEM 7 27 POTENTIAL.
FT TRANSMEM 490 510 POTENTIAL.
SQ SEQUENCE 552 AA; 62510 MW; ESDC32C858AE4D94 CRC64;
Query Match 17.2%; Score 387; DB 1; Length 552;
Best Local Similarity 28.6%; Pred. No. 2e-23;
Matches 126; Conservative 70; Mismatches 159; Indels 86; Gaps 20;
49 YGMFADAGSGTGRTHVYTFVQKMGQLPILEGEVFDG-----VKPGLSFVDPQKGAEE 102
44 YGVICDAGSGTGRFLVYVNWISTDSLELIQIEPVIYDNKPVWKISPLSTFGTKPAQAAE 103

```

```

103 TVOGLLEVAKDSIPRSHWKKTVPVLKATAGRLRLPEHKAKALLFEVKEIFRKSFFL----- 158
104 YLRPLMELAEHRHPEESKRPVTPVPIFATAGMRLIPDEQEAFL--KNLSNKLPKITSMQ 160
159 VPKGSVISMTQDEGIPAWVTNVLFTGLH-----GH-ROETVGTDLGGAET 205
161 VLKHEIRIIEGKWEIGYSIAVNYALGKFNKATLDFPGTSPAHARQKTVMIDMGGASA 220
206 QITP-LPQPEK-----TLEQTPRGVLTSEFEMFNSTYKLYTHSYLGF-----GLKAARLATLG 256
221 QIAFELPDTDFSPSSINVENINLCRDEDSLFK--YKLFVTPLFGYGVNEGIRKYEHEMLLS 278
257 ALBTEGTDGHTFRSACLPRWLEAEWIFGGVKYQYGGN--OEG-----E 297
279 KLKXQ--NGTVIQQDCCPLNLKHT-----VTLENGENFVRRTGNMTCSNEVKCLLNPE 331
298 VGPEPCVAEVLVRVVRKHLQPE--EVORGSFYASYDYDRAVDTMDIDYEGGILKVEDFE 356
332 SSSSEVCKAEAAKCYFGAVPAPSIPLSNIEMYGSEYMYSTHDV-----LGLGGQYDAENIA 387
357 RKAREVCD-----NLENFTSGSP-----FLCMDLSYITALLKDGFGFADST--VLQ 400
388 KTKQYCSKRWSTIOAESKKQLYPRADERLRTQCFKSAWITSVLHDGFS--VDKTHNKFQ 446
401 LTKKVNNIETGWLGA--TFHL 420
447 SVSTIAGQEVQWALGAMIYHM 467

RESULT 11
ENPI MOUSE STANDARD; PRT; 510 AA.
AC P55772;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Ectonucleoside triphosphate diphosphohydrolase 1 (EC 3.6.1.5)
DE (NTPDase1) (Ecto-ATP diphosphohydrolase) (ATPDase) (Lymphoid cell
DE activation antigen) (Ecto-apyrase) (CD39 antigen).
GN NTPD1 OR CD39.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95015846; PubMed=7930580;
RA Maliszewski C.R., Delespesse G.J.T., Schoenborn M.A., Armitage R.J.,
RA Fanslow W.C., Nakajima T., Baker E., Sutherland G.R., Poidexter K.,
RA Birks C., Alpert A., Friend D., Gimpel S.D., Gayle R.B. III;
RT "The CD39 lymphoid cell activation antigen. Molecular cloning and
RT structural characterization.";
RL J. Immunol. 153:3574-3583(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SvJ.
RX MEDLINE=98399871; PubMed=9710622;
RA Schoenborn M.A., Jenkins N.A., Copeland N.G., Gilbert D.J.,
RA Gayle R.B. III, Maliszewski C.R.;
RT "Gene structure and chromosome location of mouse Cd39 coding for an
RT ecto-apyrase.";
RL Cytogenet. Cell Genet. 81:287-289(1998).
CC -1- FUNCTION: In the nervous system, could hydrolyze ATP and other
CC nucleotides to regulate purinergic neurotransmission. Could also
CC be implicated in the prevention of platelet aggregation.
CC -1- Hydrolyzes ATP and ADP equally well.
CC -1- CATALYTIC ACTIVITY: ATP + 2 H(2)O = AMP + 2 phosphate.
CC -1- COFACTOR: REQUIRES CALCIUM AND MAGNESIUM (BY SIMILARITY).
CC -1- SUBUNIT: Homodimer; disulfide-linked (By similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (potential).
CC -1- SIMILARITY: BELONGS TO THE GDA1 / CD39 NTPASE FAMILY.
CC -----

```









FT	DOMAIN	499	510	CYTOPLASMIC (POTENTIAL).
FT	CARBOHYD	73	73	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	245	245	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	274	274	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	291	291	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	333	333	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	370	370	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD	457	457	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CONFLICT	203	203	G -> S (IN REF. 2).
SQ	SEQUENCE	510 AA;	57757 MW;	82F86869040D7735 CRC64;
Query Match				
Best Local Similarity				
Matches 125; Conservative 68; Mismatches 167; Indels 92; Gaps 18;				
Qy	32	GIFLSSMCPINVSASTLYGIMFDAGSTRIHVYTFVQKMPGQLPILEGVEFDSVK-PGL	90	
Db	36	GLTQNKPLPENK-----FGIVLDGSSHTSLIYIYKPAEKENDTGVSQVEECKLKGPGI	91	
Qy	91	SAFVDQPKQGAETVQGLLEVAKDSIPRSHWKKTPEVLKATAGLRLL---PEHKAKALLFE	147	
Db	92	SEPAKLGIEDIYLEACMERARTVVPKSOHAETPVYLGATAGMELLRMKNNENLASKILST	151	
Qy	148	VKEIFRKSPLVPKGSVIMTQDEGIFAWVTNVLTCQL-----HGHROETV	195	
Db	152	VAESITRYPF--DFQGARIITGOEGAYGMITINYLDDKFIQKSGWPNLKRKGDTOETY	209	
Qy	196	GTLDLGASTQITLPOFEKTLTOTPRGYLTSPEMNSTYKLYTHSVLGFGLKAARLATL	255	
Db	210	GALDLGASTQITFPVQ-NQVLE-SPENTL-HFRLYGKNYSVYTHSFLCYGKDQALLQKL	286	
Qy	256	GALTECTDQ-----HTFRSACLPRWLEAEWIFGQVKYQYCGNQEGE	297	
Db	267	-TKDLKNTNGTIHEPCFHSQVQRRMNVSHLYEAPCTREFL-TSLPPELEIQGTG----	320	
Qy	298	VGPEPCYAEVLVRVGRKLHQPEE-----VQSGSFAPSYDYDRAVDTMDIDYEGK	347	
Db	321	--FQKC-QQSIRPLFNTSYCPYSCSPDGVFLPLPQGDFAAFSAFY-----YVMG	367	
Qy	348	GILKVEDFERKARVCDNLENFTSGSPF-----LCMDLSYITALLKDG	390	
Db	368	FLNLTSEGSFQSKVTSTLEAFCS-RPWAEIQMYFGVDVKEKYLSEYCFSGTYIILTLLSG	426	
Qy	391	FGFADST--VLQLTKVNNIETGWALGATEHL	420	
Db	427	YHFTAETWKNIHFMGKVQSTSVGWTGLGYMLNL	458	

Search completed: January 29, 2004, 10:04:10  
Job time : 10.5 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 29, 2004, 09:58:54 ; Search time 16.5 Seconds  
(without alignments)  
2494.557 Million cell updates/sec

Title: US-10-091-085-7  
Perfect score: 2252  
Sequence: 1 MATSNGTFFMLVSCVCSA.....ETGWALGATFHLLQSLGISH 428  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96158682 residues  
Total number of hits satisfying chosen parameters: 283308  
Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries  
Database : PIR 76:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	616.5	27.4	479	2 T23508	hypothetical prote
2	535.5	23.8	556	2 T3109	probable guanosine
3	488.5	21.7	518	2 A40732	guanosine-diphosph
4	484	21.5	455	2 S48859	nucleoside-triphos
5	441.5	19.6	454	2 JC4616	aprase (EC 3.6.1.
6	426.5	18.9	516	2 G84442	probable nucleosid
7	398	17.7	483	2 D86276	hypothetical prote
8	382.5	17.0	485	2 T34147	hypothetical prote
9	374.5	16.6	557	2 T16696	hypothetical prote
10	371	16.5	630	2 S50463	hypothetical prote
11	366	16.3	510	2 I56242	lymphoid cell acti
12	345	15.3	572	2 T40856	probable nucleotid
13	332.5	14.8	405	2 E86376	hypothetical prote
14	279	12.4	1052	2 T04439	hypothetical prote
15	242	10.7	508	2 C86276	7A19.33 protein -
16	148	6.6	628	2 A55421	nucleoside-triphos
17	108.5	4.8	1019	2 T40813	probable cell divi
18	106.5	4.7	535	2 F97910	glucan 1,6-alpha-g
19	105.5	4.7	553	1 G1BPSV	gene 1 protein - a
20	104.5	4.6	535	2 C95040	glucan 1,6-alpha-g
21	100.5	4.5	1151	2 H71347	hypothetical prote
22	99.5	4.4	774	2 T14555	DNA polymerase hom
23	99	4.4	711	2 S66749	hypothetical prote
24	99	4.4	726	2 AB0122	probable ferric si
25	98.5	4.4	4427	2 PN0637	polyketide synthas
26	98	4.4	497	1 JC2192	subtilisin-like pr
27	96.5	4.3	287	2 T40138	probable ribosomal
28	96	4.3	477	2 JS0597	t-plasminogen acti
29	95.5	4.2	301	2 E71482	phosphatidylserine

ALIGNMENTS

RESULT 1

T23508  
hypothetical protein K08H10.4 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 17-Mar-2000  
C:Accession: T23508  
R:Gardner, A.  
submitted to the EMBL Data Library, November 1996  
A:Reference number: Z19750  
A:Accession: T23508  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-479 <WIL>  
A:Cross-references: EMBL:Z83113; PIDN: CAB05544.1; GSPDB: GN00023; CESP: K08H10.4  
A:Experimental source: clone K08H10  
C:Genetics:  
A:Gene: CESP:K08H10.4  
A:Map position: 5  
A:Introns: 36/3; 83/3; 189/1; 300/2; 412/3  
C:Superfamily: nucleoside triphosphatase chromatin-associated

Query Match	27.4%;	Score	616.5;	DB	2;	Length	479;							
Best Local Similarity	34.1%;	Pred. No.	7.4e-44;											
Matches	149;	Conservative	82;	Mismatches	167;	Indels	39;							
Gaps	16;													
QY	9	FMLVVS--CVC	SAVSHRNQQTW	FGIFLSSMCPINVS	ASTLYGIMFMDAGS	TGTRTHVYT	66							
DB	6	FSILLISFSL	SVVTTTKTY--	WCHGDGVLN----	NQHTCRFFTTIV	DAGSTGRULHYK	60							
QY	67	FVQK-----	MPGQLPIL	EGEVFDSVKPGL	SAPVDQPKQGAET	VQGLLEVAKDSIPRSH	119							
DB	61	FHIDPAIASH	GNPFK---VE	KIFQEVKPLSG	SFAKSPSADSL	EPQLQARKEVPHFM	117							
QY	120	WKKTPVVL	KATAGLRLL	PEHKAKALL	FEVKE--IFRKS	PPLVPKGSVSI	INTQDEGIGIAFW	178						
DB	118	WEKTPITL	KATAGLRLL	PCDMADDIL	ESVEERIFNS	GFPAPDAVNV	MPGSDGVSWSF	177						
QY	179	TYNFLTQ	LQH-----	GHR---QET	VTGLDLCGAST	QITFLPQPEKTL	EQTPRGYLTSPB	229						
DB	178	TLNILL	LETFTDEPT	VTGHKPAAR	SVAAAFDLGG	SGTQLTYWPNNE	AVFSEHV--CYERD	ID	236					
QY	230	MFNSTVK	LYTHSYL	GFLGKAAR	LATLGALETEG--	TDGHTFR	SACLPRLWE--	ABWIFGGV	286					
DB	237	FGHHIRL	FTHSFL	NGNLIAAR	NLIL--QLET	NEIESTHQLIT	SCMPEGYQLTEW	EY--AL	294					
QY	287	KYQYGG	NOEGBE	VGFPFC	YAEVLVRV	R--GK	LHQPBEVQ	RGSFYAFS	YYYDRAVDTMD	YVE	345			
DB	295	KF--WN	INGS	SSHSF	ESCYGT	TCKNFV	ESSEIMH	RELKGS	PVLYFS	YFFDRALNS	GLVKGN	353		
QY	346	KGGILK	VEDFER	KAREVC----	DNLEN	FTSGSP	FLCM	DLISY	TALLK	DGFG	ADSTVLQ	401		
DB	354	EGGKIEL	QFKA	AAIB	ACRREK	TEID	DDGSH	MMW	PWQCL	DLTYI	YSLR	DGYQ	FEDNQPLV	413





Db 61 NLRSLRYSVVDDGGSTGTRIHVFGY -RIESGKPVFEFRGANVASLKLHPLGSAFADDP 118  
Qy 98 KQAEETVOGLLEAVKADSPRSHWKTTPVVLKATAGLRLPEHAKALLFEVKEIFRKSPPF 157  
Db 119 DGASVSLTELVEPAKGRVPGKWETEVRLMATAGMLLELPVQEKILGVARRVLKSGGF 178  
Qy 158 LVPKGSVIMTQDQEGIFAWTVVNFPLTQGLHGRQETVGLDGLGASTQTITFL-----P 211  
Db 179 LFRDEWASVIGSGDEGVAVWVANFALGSLGGDPLKTTGIVELGASAOQVTFVSSEPMPP 238  
Qy 212 QFEXTLQTPRGYLTSEMFNSTYKLYTHSYLVGLGKAARLATLAL-----ETGTD 264  
Db 239 EFSFTI-----SFG--NVTNLYSHSLFHGQNAADKLWGLSLRDHNSAVEPTR 287  
Qy 265 GHTERSACLPR-----WLEAEWIFGVQYQGNQGEVGFPCVAEVLRVV 311  
Db 288 EKIFDPCAPKGYNDANTQXHLGSLAERSLSDSFQAGN-----YSQCSAALTIL 341  
Qy 312 RGLKHQPEEVQSGFYAFSYYYVRAVDITMI-----DYEKGGLKVEDFERKAREV 362  
Db 342 QDGNRILLIITAGPSFLFGLGEXAKWLSNMISAGERFCGEDWSK---LRVKDPSLHEEDL 398  
Qy 363 CDNLNENTSGSPFLCMLDSYITALLKQGFPGF-ADSTVLQLTKKVNNIETGHWALCA 416  
Db 399 LR-----YCFSSAYIVSLHDTGLTGLPDDERIGYANQAGDIPLDWALGA 442

RESULT 7  
D86276  
hypothetical protein F7A19.34 [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 18-Nov-2002  
A:Accession: D86276  
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;  
ansure, N.F.; Hughes, B.; Huizar, L.  
Nature 408, 816-820, 2000  
C:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A:Reference number: A86141; MUID:21016719; PMID:111330712  
A:Accession: D86276  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-483 <STO>  
A:Cross-references: GB:AB005172; NID:95080801; PIDN:AAD39311.1; GSPDB:GN00141  
C:Genetics:  
A:Map position: 1  
C:Superfamily: nucleoside triphosphatase chromatin-associated

Query Match 17.7%; Score 398; DB 2; Length 483;  
Best Local Similarity 27.4%; Pred. No. 1.9e-25;  
Matches 119; Conservative 73; Mismatches 161; Indels 82; Gaps 17;  
Qy 32 GIFLSSMCPINVSASTLYGIMFDAGSTGTRIHV--YTFVQKMPQOLPILEGEVFD-----84  
Db 59 GSLLSRRECKLR-----YSLVIDAGSGSTGRVHVFYWFESGKP-----VDFGEKH 103  
Qy 85 ----SVKPKGSFAVDQPKQAQETVOGLLEAVKADSPRSHWKTTPVVLKATAGLRLPEHK 140  
Db 104 YANLKLTPGLSSYADNEGASVTKLVEFAKQIPKMFERSDIRLMATAGRWLLEVPV 163  
Qy 141 AKALLFEVKEIFRKSPPFLPKGVSVMITQDQEGIFAWTVVNFPLTQGLHGRQETVGLD 200  
Db 164 QEQLLEVTRVLRSSGFMFRDEWANVISGSGDEGIYSWITANYALGSLGTDPLETTGIVEL 223  
Qy 201 GGASTQITFLPQEKTLTQTPRGYLTSEMFNSTYKLYTHSYLVGLGKAARLATLALGLET 260  
Db 224 GGASAOQVTFV-----SSEHPVPEYSRTIAYGNISYIYSHSFYDYGKDAALKKLEKLN 278

Qy 261 EG---TDGHTFRSACLPRLMEAEWIF--GGVKYQYG-----GNQGEVGFEPQYAE 306  
Db 279 SANSTVDG--VVEDPCTPK-----GYIDTNSKNYSSGFLADESKLQAGNFKCRSA 333  
Qy 307 VLAVVR-----CKLHQPEVQSGSFASYYYVRAVDITMIDYEKGGLKLV- 352  
Db 334 TFALLKGEKENCULYEHCSIGSTFTPD--LQGSFLATASFYITAKFFEL--BEKGWJSELI 389  
Qy 353 -----EDFERKAREVCNLENFTSGSPFLCMLDSYITALLKQGFPGA--DSTVLQLT 402  
Db 390 PAKRKYCGEWSKLLILEYPTTDEYLRG---YCFSAAYTISMHLDSLGIALDDDESITYAS 446  
Qy 403 KK-VNNIETGHWALCA 416  
Db 447 KAGEKHIPLDWALGA 461

RESULT 8  
T34147  
hypothetical protein C33H5.14 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 17-Mar-2000  
A:Accession: T34147  
R:Bradshaw, H.; Stellyes, L.  
submitted to the EMBL Data Library, November 1995  
A:Description: The sequence of C. elegans cosmid C33H5.  
A:Reference number: Z21482  
A:Accession: T34147  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-485 <BRA>  
A:Cross-references: EMBL:UA1007; PIDN:AAA82272.1; CESP:C33H5.14  
A:Experimental source: strain Bristol N2  
C:Genetics:  
A:Gene: CESP:C33H5.14  
A:Introns: 22/1; 83/1; 120/2; 167/1; 269/3; 399/3  
C:Superfamily: nucleoside triphosphatase chromatin-associated

Query Match 17.0%; Score 382.5; DB 2; Length 485;  
Best Local Similarity 27.5%; Pred. No. 3.9e-24;  
Matches 119; Conservative 84; Mismatches 153; Indels 77; Gaps 18;  
Qy 49 YGIMFDAGSTGTRIHVYTFVQKMPQOLPILEGEVFD-----VKPCLSAFVDPKQGA 102  
Db 25 YGVICDAGSGSTRFLFVYT-LKPLSGGLTNTDTLHSESPVWVKVTPGLSSFGDKPEQVVE 83  
Qy 103 TVOGLLEAVKADSPRSHWKTTPVVLKATAGLRLPEHAKALLFEVKEIFRKSPPFL-VPK 161  
Db 84 YLTPLLPFABEHIFPBLQGETDILLIPATAGMRLLPEAQKDAIINKLQGLKSVTALRVSD 143  
Qy 162 GSVMITQDQEGIFAWTVVNFPLTQGLHGRQETVGLDGLGASTQITFLPQFEK-----215  
Db 144 SNRIIDGAWEGYISWIAVNYILGRFDKENDSKVMIDMGASVQIAFEIANEKESYNGG 203  
Qy 216 TLEQTPRGYLTSEMFNSTYKLYTHSYLVGLGKAARLATLALGLETGDTGHTFRSACLPR 275  
Db 204 NYVEINLGSITNEDYK--YKIYSTTFLGVANEGLKKEYSLVKSGNS-----NDSCSPR 257  
Qy 276 WLEAEWIFGVQYQGNQGEVGFPCVAEVLRVVGRKHLQPE-----319  
Db 258 GLNR--LIG-----EFTVNGTGE--WDVCLAQVSSLI-GDKAQPSCNPCTCLERNVAPSV 308  
Qy 320 EVQSGSFYAFS--YYDRAVDITMIDYEKGGLKVEDFERKAREVC---DNLENFTSGSP 374  
Db 309 NUSTVOLYGFSEYWTTS-----NFGSGGEVHYQKFTDEVRYKIOCKDNDINDIOQGFKN 362  
Qy 375 F-----LCMLDSYITALLKQGFADST--VLQLTKKVNNIETGHWALGA-----416  
Db 363 FPNADIERLGTNCFKAAWTVSVLHDGFN-VDKTKHLFQSVLKIAGEEMQWALGAMLYHSHK 421  
Qy 417 --TFHLLQSLGIS 427  
Db 422 DLKFNLLLEQLEVA 434



RESULT 9  
T16696  
hypothetical protein R07E4.4 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Sep-1999  
C:Accession: T16696  
R:Miller, N.  
submitted to the EMBL Data Library, October 1995  
A:Description: The sequence of C. elegans cosmid R07E4.  
A:Reference number: Z18561  
A:Accession: T16696  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-557 <ML>  
A:Cross-references: EMBL:U39652; NID:G1049390; PID:G1049394; PIDN:AAA80403.1; CESP:R07E4  
A:Experimental source: strain Bristol N2  
C:Genetics:  
A:Gene: CESP:R07E4.4  
A:Introns: 39/1; 67/1; 145/2; 192/1; 244/3; 317/3; 451/3; 488/2

Query Match 16.6%; Score 374.5; DB 2; Length 557;  
Best Local Similarity 28.3%; Pred. No. 2.2e-23;  
Matches 126; Conservative 70; Mismatches 159; Indels 91; Gaps 21;

QY 49 YGIMFDAGSTGTRIHYVTFVQKMPGOLPILEGEVFDSD-----VKPGLSAFVDPQKQAE 102  
DB 44 YGVICDAGSTGTRLFVYNNWISTSDSELQIEPIYDKNPKWKISLGLSTFGTKPAQAE 103

QY 103 TVOGLLEVAKDSIPRSHWKKTTPVVLKATAGLRLLPEH-----KAKALLFEVKEIPRKSPP 157  
DB 104 YLRPLMELAEHRHPEEKRPVTPVFIFATAGMRLIPDEYVLIGQKEAVL-----KNLRNKLPK 160

QY 158 L-----VPKGSVMTGQDEGIFAWTVNFLTGLQLH-----GH-RQSTVGTLDL 200  
DB 161 ITSMQVLKEHRIIEGKWEIGYIAVYALFKFNKTATLDFPGTSPAHARQKTGVGMIDM 220

QY 201 GGASTOITF-LPOFEK-----TLQTPRGYLTSPFMENSTYKLYTHSVLGF-----GLKAAR 251  
DB 221 GGASQAIAFELPDTDSFSSINVENINLGCRDSDLFK--YKLFVTFPLGYNIGIRKYE 278

QY 252 LATLGALTEGTGHTFRSACLPRWLEAEWIFGVKYQYQGN--QEG----- 296  
DB 279 HMLLSKLKQD--NGTVIQDDCPMLNLHKT-----VTLENGENFVRRTGNWNTCSNEVKK 331

QY 297 -----EVGFPCYAEVLVRVKGKHLQPE-EVQSGSFVAFSYVYDRAVDTMDIDYKGGILK 351  
DB 332 LLNPESSEVCKAEAAKCYFCGAVPAPSIPLSNIEMYGFSYWYTHDV-----LGLGGQYD 387

QY 352 VEDFERKAREVCD-----NLENFTSGSP-----FLCWDLSYITALLKDGFGFADST 397  
DB 388 AENIAKTTQYCSKRNSTIOAESKKQIYPRADERLRTQCFKSAWITSVLHDGFS-VDKT 446

QY 398 --VLQLTKYNNIETGHWALGA-TFHL 420  
DB 447 HNKFSQSVTIAGQEVQVQALGAMIVHM 472

RESULT 10  
S50463  
hypothetical protein YER005w - Yeast (Saccharomyces cerevisiae)  
C:Species: Saccharomyces cerevisiae  
C:Date: 28-May-1993 #sequence\_revision 24-Feb-1995 #text\_change 19-Apr-2002  
C:Accession: S50463  
R:Dietrich, F.S.  
submitted to the EMBL Data Library, December 1994  
A:Description: The sequence of S. cerevisiae cosmid 9537, 9581, 9495, 9867, and lambda  
A:Reference number: S50433  
A:Accession: S50463  
A:Molecule type: DNA  
A:Residues: 1-630 <DIE>  
A:Cross-references: EMBL:U18778; NID:G603592; PID:G603597; GSPDB:GN00005; MIPS:YER005w

C:Genetics:  
A:Gene: SGD:YND1; MIPS:YER005w  
A:Cross-references: SGD:S0000807  
A:Map position: 5R

Query Match 16.5%; Score 371; DB 2; Length 630;  
Best Local Similarity 26.0%; Pred. No. 5.3e-23;  
Matches 116; Conservative 75; Mismatches 164; Indels 92; Gaps 17;

QY 49 YGIMFDAGSTGTRIHYVTF-----VQKMPGOLPILEGE-----VFDVSKPGLSA 92  
DB 10 FGIVIDAGSSGRIHVFKWQDTESLLHATNQDSQILQSVPHIQEKDWTFLKLPGLSS 68

QY 93 FVDQPKQGAET-VQGLLEVAKDSIPRSHWKKTTPVVLKATAGLRLLPEHAKALLFEV-KE 150  
DB 69 PEKKPQDAYKSHIKPLLDPAKNIIPESHWSCEVFIQATAGMRLLPQDIQSSSLDGLCOG 128

QY 151 IFRKSPFLVP--KGSVMTGQDEGIFAWTVNFLTGLQLHGRQE-----TVGTLDLGA 203  
DB 129 LKHPAEFLVEDCSAQIQVIGETGLYGLWGLNYLGHFNFDYNPEVSDHFTFGFMDMGA 188

QY 204 STQITFLPQFEKTL-----OTPRGYLTSPFMENSTYKLYTHSVLGFGLKAARL 252  
DB 189 STQIAFAPHSGEIAHRDDIATIFLRSVNGDLQKWDVFSV-----WLGFGANQARR 241

QY 253 ATLGA-----ETEGTDGHTFRSACLPRWLEAEWIFGVKYQYQGNQGEVGFPE 301  
DB 242 RYLAQLINTLPENTDYNDFSTRNLNDPCMPRGSSSTDFEFKDTIFHLAGSGNYEQCTK 301

QY 302 PCYAEVLR-----VVRGKHLQPE-EVQSGSFVAFSYVYDRAVDTMDIDYKGGILK 352  
DB 302 SIYPLLLKMPDCDBPCLFNGVHAPRIDFANDKFTGTSEYWTYANDV-----PKLGGEYNF 357

QY 353 EDEPKAREVCN-----LENFTSG-----SPFL--CMWLSYITALLKDGFGFA--- 394  
DB 358 DKFSLRLSEFCNSNWTQILANDSKGVNSIPENFLKDACFKGNWVNLHLEGFDMRIDV 417

QY 395 -----DSTVLQLTKYNNIETGHWALG 415  
DB 418 DAENVNDRPLFQSVKVEERELSWTLG 444

RESULT 11  
I56242  
lymphoid cell activation antigen - human  
C:Species: Homo sapiens (man)  
C:Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 18-Nov-2002  
C:Accession: I56242  
J:Maliszewski, C.R.; Delepesse, G.J.; Schoenborn, M.A.; Armitage, R.J.; Panselew, W.C.;  
R:Immunol. 153, 3574-3583, 1994  
A:Title: The CD39 lymphoid cell activation antigen. Molecular cloning and structural cha:  
A:Reference number: I56242; MUID:95015846; PMID:7930580  
A:Accession: I56242  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-510 <RES>  
A:Cross-references: GB:S73813; NID:G765255; PIDN:AA32152.1; PID:G765256  
C:Superfamily: nucleoside triphosphatase chromatin-associated

Query Match 16.3%; Score 366; DB 2; Length 510;  
Best Local Similarity 28.2%; Pred. No. 1e-22;  
Matches 125; Conservative 72; Mismatches 172; Indels 74; Gaps 19;

QY 32 GIFLLSMCPINYSASTLYGIMFDAGSTGTRIHYVTFVQKMPGOLPILEGEVFDVSK-PGL 90  
DB 36 GLTONKALPENVK-----YGLVDAGSSHTSLYIKWPAEKENDTGVVHQVEECRVKPGI 91

QY 91 SAFVDQPKGAETVQGLLEVAKDSIPRSHWKKTTPVVLKATAGLRLL---PEHKAKALLFE 147  
DB 92 SKFVQKVEIGIYLTDCMERAREVIPSQHOETFPVYLGATAGMRLLRMESEELADRLDV 151

QY 148 VKEIFRKSPLVPKGSVMTGQDEGIFAWTVNFLTGLQ-----HGRQETV 195

Db 152 VERSLSNVPF--DFQARIITQBEAGYGVMTITVNLGKFSQKTRWRSIVPYETNNQETF 209  
Qy 196 GTDLGGASTQITELPOEKTLEOTPRGYLTSFEMFNSTYKLYTHSYLGFGLKAARLATL 255  
Db 210 GALLDGGASTQVTFVQ--NQIIE--SPDNAL--QFRLYKDYNNVYTHSFLCYGKQALWQKL 266  
Qy 256 GALTETGTDGHTFRSACL-----PRWLEAEWIFGGVKYQYQGNQGE 297  
Db 267 -AKDIQVASNEILRDCPHGPKYKVVNSVDLYKTPCTKRFEMTLPPFOFEIQIGN--- 321  
Qy 298 VGFPEPCVAELRV-----VRCKLHQPVEQVQSGFYASYYDYDRAVDTMDIYE 345  
Db 322 --YQCHQSILELFTSYCPYSQCAFNGIFLPLQDGFCAFSAF--YFVMKFLN---LTSE 375  
Qy 346 KGGILKVEDFERK-AREVCNLENFTSG--SPFL---CMDLSYITALLKDGFGF-ADS-T 397  
Db 376 KVSQEKTEMKKCAQWBEIKTSYAGVKEKYLSEYCFSGTYILSLILLOGHFTADSWE 435  
Qy 398 VLQTKKNNIETGALGATPHL 420  
Db 436 HIHFIGKIQGSDAGWTGLGYMLNL 458.

RESULT 12  
T40856  
probable nucleotide phosphatase - fission yeast (Schizosaccharomyces pombe)  
C:Species: Schizosaccharomyces pombe  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999  
C:Accession: T40856  
R:Ramsperger, U.; Pohl, T.; Wood, V.; Rajandream, M.A.; Barrell, B.G.  
submitted to the EMBL Data Library, October 1999  
A:Reference number: Z21952  
A:Accession: T40856  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-572 <RAM>  
A:Cross-references: EMBL:AL121783; PIDN:CAB57847.1; GSPDB:GN00068; SPDB:SPCC11E10.05c  
A:Experimental source: strain 972h-; cosmid c11E10  
C:Genetics:  
A:Gene: SPDB:SPCC11E10.05c  
A:Map position: 3

Query Match 15.3%; Score 345; DB 2; Length 572;  
Best Local Similarity 25.3%; Pred. No. 7.1e-21;  
Matches 114; Conservative 69; Mismatches 156; Indels 112; Gaps 16;  
Qy 49 YGIMFDAGSTGTRHVTFF-----VQKMPGQLPILEGEVFD-----SVKGLSAFVD 95  
Db 5 YGIFIDAGSSGRLIYSDYDTSLSLSDKVKKPLLIETGIGDGKWSLKVQPGISSFAN 64  
Qy 96 QPKQ-GAETVQGLLEVAKDSIPRSHWKTTPVVLKATAGLRLL-PEHKAKALLFEVKEIFR 153  
Db 65 NPKHVGHKHLKELLDFPAHAIPKDVHKTETPVLSATAGMLLGVDAQNKILSHACRYIKK 124  
Qy 154 KSPFLVP--KGSVIMTQDEGIFAWTVNFLTQGLHGRQETVGTLDLGGASTQITFLP 211  
Db 125 NYDFDIPNCNSIRVIDGKAEGMYGLATNYLLKTEEDKSTSTVGFLDMGGASVQIAF-- 182  
Qy 212 QFEKLTQTPRGYLTSEMFSNT-----YKLYTHSYLGFGLKAARLATLIGAL- 258  
Db 183 -----ELPPSQKLNKYSISTVHIGLQNGQQQLEYFLPVTWLGFGANEAYRYLGLLI 235  
Qy 259 ETE-GTDGHTFRSACLPRWLEAEWIFGGVKYQYGG-----NOEGEVGPEPCY 304  
Db 236 ESENGKVGNTLSDPCSRL--GRTYDIDGIEFAGTGDILKQCLKLTYNLLNKDKPCSDMPCN 293  
Qy 305 AEVLVRVGRKLHQPVEQVQSGFYASYYDYDRAVDTMDIYKGGILKVEDFERKAREVC- 363  
Db 294 FDGISI-----PPVDPAANTEFVGVSFEWTTNDV-----FDMGGSYHFPNFKYKVDYCYG 343  
Qy 364 -----DNLNFTSGSPFLCMDLSYITALLKDGFGFADSTV----- 398  
Db 344 TEWETMLSRNLNKLTPSTDENKLEK-----LCFKASWALNVLHGFDPVPSKNTSSND 396

Qy 399 -----LQTKKVNNIETGWAIG 415  
Db 397 AKDGLSVIPAYHSPFTSLEKIERTEVSWTIG 427

## RESULT 13

E86276

hypothetical protein F14117.1 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Dec-2001

C:Accession: E86276

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.  
ansen, N.F.; Hughes, B.; Huizar, L.  
Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.;  
C.A.; Li, J.H.; Li, Y.; Lin, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: E86276

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-405 &lt;STO&gt;

A:Cross-references: GB:AE005172; NID:g7262666; PIDN:AAF43924.1; GSPDB:GN00141

C:Genetics:

A:Map position: 1

Query Match 14.8%; Score 332.5; DB 2; Length 405;  
Best Local Similarity 26.3%; Pred. No. 4.9e-20;  
Matches 110; Conservative 64; Mismatches 157; Indels 87; Gaps 15;

Qy 48 LYGIMFDAGSTGTRHVTFFVQKMPGQLPILEGEVFDVSKVGLSAFVQPKQGAETVQGL 107  
Db 3 VFGYWFESGKP-----VFDGEEHYASL-----KLSPGLSSYADNPEGASVSVTKL 48

Qy 108 LEVAKDSIPRSHWKTTPVVLKATAGLRLLPEHKAKALLFEVKEIFRSPFLVPKGSVIM 167  
Db 49 VEFAGRIKPKLKKSDIRLMATAGMRLLDVPVQEQILDVTRVRSSGPKFQDEWATVI 108

Qy 168 TGQDEGIFAWTVNFLTQGLHGRQETVGTLDLGGASTQITFLPOEKTLEQTPRGYLT 227  
Db 109 SGTDGEGYAVVANHAGLSLGGDPLKTTGIVELGGASQVTFVPS-----EHPPEFSRT 163

Qy 228 FEMFNSTYKLYTHSYLGFGLKAARLATLGALETE---GTDGHTFRSACLPR----- 275  
Db 164 ISYGNVSYTYSHSFLDFGQDAADKLLSLSQNSVAASGTDGIVDPCTPKGIYDTHSQ 223

Qy 276 -----WLEAEWTF-GGVKYQYQGNQGEVGFEPFYAEVLVRVR-GKLHQPVEVQSGFYA 328  
Db 224 KSSGFLSBESKPKASLQVQAAGD-----FTKCRSATLMLQEGKEN----- 265

Qy 329 FSYVDRAVDTMDIYKGGILKVEDF-----ERKAREVCNLL 366  
Db 266 -CAYKHCISGISTPTNIOGSLATENFPHTSKFFGLGEKWLSEMLAGKRCFGEWSKL 324

Qy 367 -ENF-TSGSPFL---CMDLSYITALLKDGFGFA-DSTVLQLTKKV--NNIETGWAIGA 416  
Db 325 KERYPPTKDKYLHRYCFSSAYIISMLHSLGVALDDIRIKYASKAGKAGINIPLDWALGA 382

## RESULT 14

T04439

hypothetical protein T18B16.150 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 30-Apr-1999 #sequence\_revision 30-Apr-1999 #text\_change 30-Apr-1999

C:Accession: T04439

R:Bevan, M.; Benes, V.; Rechmann, S.; Borkova, D.; Ansoorge, W.; Bancroft, I.; Mewes, H.W.  
submitted to the Protein Sequence Database, April 1998

A:Reference number: Z15359



**THIS PAGE BLANK (USPTO,**